DOCTOR OF PHILOSOPHY

Developing skills to explain scientific concepts during initial teacher education the role of peer assessment

Cabello Gonzalez, Valeria Magally

Award date:
2013

DOCTOR OF PHILOSOPHY

Developing skills to explain scientific concepts during initial teacher education

the role of peer assessment

Valeria Magally Cabello Gonzalez

2013

University of Dundee

Developing skills to explain scientific concepts during initial teacher education: The role of peer assessment

Valeria Magaly Cabello González

Thesis accepted for the degree of Philosophy Doctor (Ph.D.)

in Educational Psychology

May, 2013

Contents

List of Tables

List of Figures

Acknowledgements

Over the past years I have received contributions from a large number of admirable people, who have kept my motivation high to complete this research and thesis. Professor Keith Topping was my first supervisor and he helped me from the very beginning to transform my initial ideas into a complete study. He offered me wise guidelines, remarks, fruitful recommendations and assertive corrections to name a few of his contributions. I would like also to mention in this first place Professor Susan Rodrigues, who was my second supervisor during the first year of my Ph.D. studies. She shared with me her generosity, time, advices, references and illuminating ideas. Likewise, Professor Norman Reid took an important role on the supervision of my work during the two last years, giving me insightful inputs to challenge my ideas. Also Neil Taylor joined my advisory team in order to help my understanding of science teaching in real teaching contexts, and to provide me all extra support, including enthusiasm and encouragement when I needed them. Thanks to all.

This research required the participation of three universities. I am grateful to all the people who not only gave me access to their institutions, but also valuable discussions about the project. In particular, I would like to thank Dr. Luis Eduardo Ravanal and David Santibáñez who generously worked with me during data collection. I would also like to thank all the participants of this study that believed in my skills to make their undergraduate education more meaningful and effective. My ex colleagues from Proyecto Docentemás provided concrete material and good ideas for the validation of the main instrument used in this research. Additionally, Cristian Arnés, Paula Cavada, Paulina Ramirez and Andrés Andrades helped this project to reach its validity and reliability. Similarly, I want to thank the Chilean students who demonstrated during 2011 their concerns about the quality and equity of our educational system. They were a source of inspiration and challenge to work for a better education. Besides, the Chilean government provided financial support for the study.

Finally I would like to extend my deep gratitude to my family for supporting my work from far away: my parents Jorge Cabello and Magaly González, my siblings, grandparents and my dearest Clément Demons. They gave me unwavering love and reinforcement to complete this research. My friend Ximena Poblete also encouraged me since this project was just an illusion. Thank you all for believing this doctoral project was possible despite the difficulties.

Declaration

Valeria Magaly Cabello González is the author of the thesis titled “Developing skills to explain scientific concepts during initial teacher education: The role of peer assessment”.

All the material cited in this thesis have been properly quoted or attributed. All the references cited have been consulted by Valeria Magaly Cabello González and their source can be accessed in the list at the end of this document.

The work of which this thesis is a record has been done by Valeria Magaly Cabello González and it has not been previously accepted for a higher degree. It was conducted following the requirements of the University of Dundee Research Ethics Committee (University of Dundee, 2007), body that approved the research design on 24th May 2011: Application Number 11031.

Valeria Magaly Cabello González

Abstract

Initial teacher education is an area of weakness within the Chilean education system. Yet it is highlighted as a crucial aspect of educational success. Success in educational improvement depends mainly on the teachers (because they enact a reform by putting it into practice), and teacher thinking is likely to influence teacher decision-making. How teacher conceptions and practice change, and how to facilitate this change, was the focus of this study. It explored to what extent peer assessment could facilitate change in pre-service science teachers' conceptions and practices regarding conceptual explanations in science teaching.

In a quasi-experimental design, a ten-session peer assessment intervention was carried out with thirty seven pre-service science teachers in three Chilean universities, each with an experimental and control group. The intervention sought to develop changes in teachers' conceptions about the quality of explanations and in their skill of explaining scientific concepts. Teachers' thoughts were obtained through a peer assessment questionnaire, feedback sessions, focus groups and interviews. The quality of their explanations was measured at pre, post and follow-up in their eventual first job via video-recorded microteaching episodes using observational analysis. Inter-rater reliability was calculated on 5% of all qualitative data and all the videos were rated by two researchers in a blind process. Qualitative analysis indicated how teachers transformed their conceptions about the quality of explanations from general pedagogical knowledge into pedagogical content knowledge. A quantitative instrument was created to evaluate student teachers' explanations in practice. Its reliability enables the assessment the skill of explaining based on ten elements (Cronbach's alpha=.77). Results showed pre-service teachers significantly improved their explanations of scientific concepts in some practical aspects, although not all of them were transferred into real teaching contexts. The changes in student teachers' conceptions and practice were analysed to indicate how the process occurred, to what extent peer assessment had a role on it, and which elements facilitated or made difficult the transference of the skill of explaining into real teaching. These results indicated that peer assessment can play a noteworthy role in teacher education to develop skills. There are implications for policy and practice in this study, not only for teacher education but also for in-service teacher professional development, not only for Chile but also for other countries.

List of Abbreviations and Acronyms

The following list presents the meaning of the abbreviations and acronyms used throughout the present thesis. The page on which each one is defined or first used is also given.

Abbreviation	Meaning	Page
ITE	Initial Teacher Education	14
ERIC	Education Resources Information Centre	15
PA	Peer Assessment	16
OECD	Organisation for Economic Co-operation and Development	16
PISA	Programme for International Student Assessment	16
TIMMS	Trends in International Mathematics and Science Study	23
SIMCE	Sistema de Medición de la Calidad de la Educación (Quality of Education Measurement System)	25
OREALC-UNESCO	Oficina Regional de Educación para América Latina y el Caribe (Regional Bureau of Education for Latin America and the Caribbean). United Nations Educational, Scientific and Cultural Organization	27
NSTA	National Science Teachers Association (U.S.A.)	28
ASTA	Australian Science Teachers Association	28
GTE	General Teaching Council of Scotland	28
QAA	Quality Assurance Agency for Higher Education	29
PK	Pedagogical Knowledge	30
CK	Content Knowledge	30
PCK	Pedagogical Content Knowledge	30
ICT	Information and Communication Technology	43
U1, U2, U3	University 1, University 2, University 3	85
PSU	Prueba de Selección Universitaria (University Selection Test)	89
SPSS	Statistical Package for the Social Sciences	101
QSR	Qualitative Solutions and Research	101
SQ	Scored Quality criteria of the explanation	107
GA	General Aspects	113
PKA	Pedagogical Knowledge Aspects	113
PCKA	Pedagogical Content Knowledge Aspects	113
KA	Knowledge Aspects	113
T1, T2… T20	Teacher 1, Teacher 2… Teacher 20	130
F1, F2, F3	Focus Group 1, Focus Group 2, Focus Group 3	133
I1, I2…I6	Interview 1, Interview 2… Interview 6	145

1. Introduction

1.1. Rationale

There is much public international discussion about the need to develop teacher quality and teachers' knowledge (Lawson, Askell-Williams, & Murray-Harvey, 2009). In this debate, it has been suggested that to improve the quality of teacher education programmes, these programmes should incorporate different types of assessment and accountability to encourage teachers to self-reflect on their teaching and make the necessary adjustments to develop their teaching (Borman, Mueninghoff, Cotner, & Frederick, 2009).

Some of the key questions posed in initial teacher education (ITE) are about the transition from being a student teacher to a beginning teacher (when they begin to teach in schools): How does the student teacher transform subject matter into a form that pupils can comprehend? When the beginning teacher confronts flawed or muddled content in textbook chapters or confused students, how does the teacher generate new explanations, representations or clarifications? What are the sources for analogies, metaphors, examples and demonstrations? (Shulman, 1987). These questions inspired this work. Specifically, according to Trout (2002, p. 212) few products of intellectual life are more exhilarating, more pleasing to give and receive, than a good explanation. In his words, a good explanation "feels right". Although the occurrence of this sense or feeling of understanding is neither necessary nor sufficient for a good explanation, it does drive judgements of the plausibility and ultimately, the acceptability of an explanation. Philosophers of science have dealt extensively with the topic of explanation and understanding, and several questions have also arisen for science education. For instance, do science educators have clear notions and criteria for explaining science? (Edgington, 1997).

Nevertheless, the number of studies about the nature of teachers' explanations is insufficient to provide a guiding framework to understand the explanations in science education as an object of study (Dagher & Cossman, 1992; Geelan, 2012). According to these researchers, most of the research into the topic of explanations has been centred on students' explanations in science (Edgington, 1995; Mestre, Dufresne, Gerace, Hardiman, & Touger, 1993; Sevian & Gonsalves, 2008; Tamir & Zohar, 1991) and not on the science teacher's explanations necessarily. This diversity indicates that further research will enrich the area of science education, because it might contribute to develop a research framework for the explanations of natural phenomena, serving to analyse and guide teaching work in different disciplines (Edgington, 1997; Geelan, 2009). The most representative research about science teacher explanations were conducted during the nineties' (Dagher, 1992; Treagust & Harrison, 1999).

Nowadays the situation does not differ much from the one described above. A small number of research projects has been done into science teachers' explanations, in spite of its importance in science education (Geelan, 2009, 2012). Actually a meta-analysis recently reported by Geelan (2012) indicated that researching in Education Resources Information Centre (ERIC) data base with the terms 'science teach* explain*' retrieved 1362 hits, but from these fewer than 35 papers were focused on an aspect of teacher explanations in science. Then, the potential in the area of research about explanations has yet to be recognised (Edgington, 1997). The studies have been sparse to date and there remains the scope for much more research to be done (Geelan, 2012). Actually, science teacher explanation area has been stated as a fruitful field for new research because it has a potential to offer considerable new insight into science teaching and learning, and it can lead to vast contributions to science teacher education (Geelan, 2009, 2012). Furthermore, Dagher and Cossman (1992, p. 362) asserted that "we need to inquire about teachers' explanatory behaviour". Similarly, Geelan (2012) concluded that the field of research about teacher explanations must be developed, seeking to improve the quality of the explanations given by teachers.

Similarly, Sampson and Clark (2007) mentioned that teacher explanations could serve as models for pupils, as they may learn to explain and also to argue about their own scientific ideas from teacher's explanations. This is another good reason why teacher explanations need to be addressed and developed during initial teacher education (Geelan, 2012), which is concordant with the ideas of Sevian and Gonsalves (2008), who stated the need of training for those who teach science to university students in higher education to construct effective explanations of science. Constructing explanations is a task that requires abstract thought (Cobern & Loving, 2000), and explaining abstract and complex concepts in everyday terms could be considered as a test of the explainer's understanding (Feynman, 1994). Even though most individuals have had exposure to school teachers and have some perception of what is a good teacher or teaching (Murphy, Delli, & Edwards, 2004), there is not a reported consensus in terms of what constitutes a good science teacher explanation or how to assess its quality.

In the field of educational studies, besides, there is an increased interest in assessment for learning, which can be considered one of the purposes of peer assessment (Gielen, Dochy, & Onghena, 2010). Peer assessment (PA) as defined for this thesis, is understood as the agreement between peers to consider the quality or successfulness of the teaching performance from other similar status learners, in this case, peer pre-service science teachers. There are just a few studies conducted so far in science teacher education using PA, especially formative PA. This is considered an issue because teachers are exposed to the observation and assessment of their work by peers during their professional lives, which usually carries negative feelings and the resulting resistance to cooperative work (Kukanja, 2007). As a result, it was interesting to ask how formative PA of the pre-service teachers' explanations would work in science ITE.

Furthermore, as the teaching practice is one of the most important aspects of teacher education (Oluwatayo & Adebule, 2012) but in the context of Chile it is treated as one of the least important (Vergara & Cofré, 2008), it was decided to offer pre-service teachers the possibility to perform and assess their explanations of scientific concepts in the present research. This teaching practice was settled through simulated microteaching episodes, during pre-service teachers' final year of undergraduate education. In the context of Chile, focusing on the formative intention to improve science teacher explanations was even more important, because it has been demonstrated that the most frequent strategy that teachers use in science classrooms is the conceptual explanation (Preiss, Alegría, Espinoza, Núñez, & Ponce, 2012).

In Chile, in-service science teachers are older than in the Organisation for Economic Co-operation and Development (OECD) countries, reported by the Programme for International Student Assessment (PISA) 2006 (Gobierno de Chile, 2007). Besides, they have a limited or no specialization in science teaching (Cofré et al., 2010). Consequently, teachers feel less secure to teach in areas like chemistry, physics and earth sciences than teachers in other OECD countries (OECD, 2006). It is important to notice that the term "science teachers" is being used in this research as it is defined in the context of Chile: teachers of the three basic sciences that teach at secondary level (ages 15-18) and generalist teachers with specialization in general sciences, who teach at primary level (ages 11-14). Nowadays there is a 35% shortage of science teachers (Palma, 2012) which is a source of concern for the government to attract and maintain science teachers teaching in schools.

Furthermore, science teacher education in Chile is in a complex situation due to the lack of connection between teaching practice and the enormous variety of teacher education programmes that have the title of science teacher education with no minimum standard or common competencies established (Cofré et al., 2010). In this sense, this project is coherent with the need of exploring teacher education programmes that have a weaker practical component than already studied European programmes (Vermunt & Endedijk, 2011).

1.2. Purpose

The general purpose of this research was to explore the question: to what extent PA could facilitate change in Chilean pre-service science teachers' conceptions and practices to explain scientific concepts? Specifically, this study sought to explore the implicit theories about the quality of conceptual explanations in pre-service science teachers to determine whether differences existed according to pre-service teachers' science knowledge. The science knowledge was determined by the number of science courses offered by three participant universities to their student teachers (fourteen, nine and four). Also, this research analysed and compared the conceptions about the quality of teacher explanations of pre-service science teachers in experimental group vs. a control group. Besides, the quality of pre-service science teachers' conceptual explanations was identified before and after PA in the experimental group, and the reasons of possible changes were collected from the teachers' and researcher's perspective.

Barnett and Hodson (2001) argued that teachers must be able to generalise some aspects of knowledge and skills to new situations. From this, an interesting question was how generalizable and transferable good practice to explain was? Thus, to look for the generalizability and transference of the possible improvements gained during PA into a real teaching context, this research compared student teachers' practice during PA and in their first job in real schools, identifying also the facilitators and obstacles for the transference.

This last stage was considered crucial because nowadays research into science teaching is approaching a new phase whereby 'testing for applicability' in classroom practice might encourage new conceptualizations of the perceived purpose of educational research (Loughran, Berry, & Mulhall, 2007). Likewise, exploring teachers' thoughts at the beginning and at the end of PA was based on the idea of encouraging them to look at themselves reflectively in order to bring about a transformation of their views (Ferguson, 2008).

The instruments created in this research intended not only to assess but also to trigger development, following the conceptualization of Avalos (2011) in terms of adapting the instruments to the project objectives but also to teachers' needs in order to help their professional development. From her perspective, "professional development is about teachers' learning, learning how to learn and transforming their knowledge into practice for the benefit of their students' growth" (Avalos, 2011, p. 10).

1.3. Paradigm

According to Gielen et al. (2010) it is important to define in PA studies -such as the present research- the objectives on which PA was based. Here, , PA was oriented to collective construction of meaning, the application of student teachers' own theories about the quality of explanations in science to their current skills to explain science, in order to develop critical thinking and self-reflection. The paradigm that supported this work was social constructivism, in this case, applied to science teachers and teaching processes (Fenshamp, Gunstone, & White, 1994; Tabachnick & Zeichner, 1999). It is important to mention that language and teachers' speech were considered in this research as tools and products of cognitive, social and cultural practice following the ideas of Vygotsky presented by Cole, John-Steiner, Scribner, and Souberman (1978). Also, understanding in the sciences are considered as constructed within social and cultural interactions and scientific knowledge production as maintained by social interchange in the classroom (Moje, Collazo, Carrillo, & Marx, 2001).

Consequently, the implementation of PA intervention was underpinned on action research. In action research according to Sandoval (2002) the process of interaction is in the origin of social reality that is endowed of meaning through what participants do during their action.

In this research it took the form of a practical seminar to de-construct and co-construct science teachers' conceptions and practice about explanations of scientific concepts in the classroom. In the perspective of action research, action is considered as a valuable tool to promote systematic processes of development (Lebak & Tinsley, 2010; Tabachnick & Zeichner, 1999). From this perspective, the participants created understanding from their social and material reality, and the participation of the researcher in the process was seen as a methodological resource to achieve the expected results. This was decided taking the social constructivist paradigm to understand and interpret how knowledge is transformed by groups of people. This paradigm according to Sandoval (2002) assumes that the comprehension of the human phenomena being investigated is a shared creation between the research participants' and the researcher's interpretations.

The intervention's perceived effectiveness was evaluated allowing participant teachers to express their thoughts and/or concerns about the intervention, considering their perspective as a source of understanding. From the point of view of LeCompte (2000), tacit or implicit theories guide daily behaviour, sometimes explaining past behaviours and predict what will happen next. Also they guide teachers' ideas. To study student teachers' conceptions about the quality of explanations was necessary to de-construct and re-construct the theories underpinning teachers' ideas. This reaffirms the interpretative paradigm this research had.

1.4. Significance

As mentioned, an important gap in the review of literature was found regarding the usage of formative PA in science teacher education. Considering this, merely by implementing the present action research in this unreported field is an original contribution to the scope of applicability of PA. In this development, this research provided inputs about how PA works in assessment of performances such as in teaching during microteaching episodes. Specifically, the researcher interpreted from teachers' comments and the position they adopted in their discourse when giving feedback to their peers, that two mechanisms were having a role in the application of PA as a facilitator of teachers' change process: the projection and reflection. This finding enriches the understanding of the underpinning principles that could make peer feedback and assessment a formative tool, and presents an original contribution to theory.

Likewise, in the review of literature there was a gap about science teachers' explanations, specifically in the criteria of quality that could make an explanation of better or worse. In that context, the present research has created a rubric to evaluate the quality of science teachers' explanations and this may prove valuable as an original contribution. This involved ten observable criteria in different achievement levels, useful to diagnose and intervene pre-service science teachers' skills to explain scientific concepts in ITE or continuous development. At the same time, this research presented a definition of quality of science teacher explanations modelled by the ten criteria mentioned, and a statistically reliable way to measure it in teachers' practice.

Also, another relevant gap in the review of literature was about the generalisation or transference of skills gaining from PA to broader or different contexts. Usually the studies in PA do not include follow-up of long term impact and this was stated as a need for future research (Sluijsmans, Brand-Gruwel, van Merriënboer, & Bastiaens, 2002). Similarly, Pauline (1993) asserted that the main problem of simulated microteaching is its difference from classroom settings, because the skills gained there would be difficult to transfer. However, in this research, this gap was filled by investigating the transferability of the skill to explain scientific concepts into real teaching. Results of the follow-up study showed that the participants not only generalised the skills acquired during PA of microteaching episodes to other contexts, but also maintained eight of the ten practical aspects development. Thus, the present research identified which aspects can be transferred and which others need more work.

Moreover, despite the presence of explanations in every form of teaching, until the year 2000, little attention was paid to the role of teachers' pedagogical content knowledge and teacher explanations according to Treagust and Harrison (1999). In this sense, this research introduced some clues about how both can be developed together.

The prior contributions to knowledge are relevant within a research area scarcely reported so far. Also, because even though the context of this study was only one country, the universities' characteristics permit its applicability to other similar countries. Likewise, the findings of this research can be useful to orient countries that currently have standards for teacher education, because it presented an original proposal to measure a competence. Thus, implications of this research are not limited to Chilean boundaries but go beyond this context.

Finally, considering the multiplicative effect that teachers' practices can have, it is possible to think that this improvement in teachers' practice might lead to better outcomes for their pupils. Following the circle, when pupils have a low achievement in science during the school, it is passed as a low scientific comprehension in adulthood, and often in a detachment of the science and technology present in the world nowadays (Frisch, Camerini, Diviani, & Schulz, 2011). Therefore, when teachers provide good science teaching, students can make better everyday life decisions and also feel more interested in developing themselves in the scientific or technological professions (Bencze & Bowen, 2009). In this sense, the current study opens the potential research field on science teacher explanations, because "classroom explanations cannot rest solely on the quality of the product" (Treagust & Harrison, 1999, p. 32). This idea may imply assessing the teacher explanations including the learning outcomes, which is one of the several projections of this study for further research.

1.5. Thesis structure

Chapter 1 is the introduction to the research and the present document. Following this, in Chapter 2 the literature review presents the main theoretical understanding and empirical studies which are relevant to this research. In Chapter 3, the methodology is described, comprising of the detailed objectives and research questions, the PA design, measurements, instruments, procedure and the data analysis techniques. Chapter 4 presents the main results of this research, organised into three studies: (1) construction of an instrument to assess quality of teacher explanations (2) PA intervention (3) Follow-up of the quality of teachers' explanations. Chapter 5 presents a critical discussion of the methodology used and the obtained results from the perspective of literature review and other relevant topics. Finally, Chapter 6 introduces the most important conclusions of this piece of research. In this chapter, actions for future research and implications for policy and practice are also presented.

2. Literature review

In this chapter an overview of the relevant topics, theories and empirical studies for this research are presented. First, a comparison between important countries to which Chile relates in terms of its aspirations in science education is introduced. The structure of science education, learning outcomes and the main factors empirically associated with those outcomes are described as well.

Improving science education has been internationally recognised as an educational goal. In the second section of this literature review the main policies and strategies that in several countries including Chile have been adopted to promote science education are presented. These involve changes in the curriculum of science education in schools and the creation of standards for science teachers. The standards are usually divided into pedagogical methods and content standards, but there is a gap between both that is filled with pedagogical content knowledge. This is a special kind of teacher knowledge and its development in teacher education is discussed as another important strategy.

The third section of this chapter concerns the current trends in teacher education, the views that have influenced the programmes, their structures and challenges that characterise the formation process of becoming a science teacher. One of the especially interesting ideas for this study area in teacher education is presented in more detail; connecting theories and practice of science pre-service teachers.

In the fourth section peer assessment in teacher education is discussed. Nowadays this type of assessment is attracting more interest among teacher educators and researchers. Its definition, application and some studies close to the current research are identified. Peer assessment has been recognised as a powerful tool to promote teachers' critical thinking. Thus in the fifth section there is a review of teachers' beliefs or conceptions, implicit theories and remarks about their modifiability as part of teachers´ thinking.

The literature review ends with a conceptual and empirical review on the focus of this research: science teachers' explanations. Its definition, differentiation from scientific explanation, characterization and issues to identify the quality of explanations in science teaching are indicated in the sixth section. This is in order to give the reader a complete view about the general and specific background in which this research was situated.

2.1. Science education and learning outcomes

Although there is agreement about the idea that science education is important for all children, there has been little debate about its nature and structure (Osborn & Dillon, 2008). Science education is across the world a priority nowadays, but it seems to be more important in some countries than in others (OECD, 2010a). A brief review of some of those countries which are of interest for the context of this research is given in the following pages, including science education structure, student achievements in science and factors associated with their learning outcomes.

2.1.1. Europe and the United States of America

a. Structure. According to Osborn and Dillon (2008) the structure of science education in Europe is based on the foundational knowledge of the three main areas of science: biology, chemistry and physics. The goal of science education is to offer education that develops all students' understanding both of the canons of scientific knowledge and how science functions, not only focused on those that will follow science careers. This is very similar to the view that the United States already have held since two decades (Rutherford & Ahlgren, 1990).

b. Student achievement in science. Although the results of most of the countries in Europe are around or above the average of OECD countries in international tests like PISA or Trends in International Mathematics and Science Study (TIMMS) (OECD, 2010b), there is a strong negative correlation between students' interest in science as a further area of study and their achievement in science tests (Osborn & Dillon, 2008). A similar trend is reported in the United States, where despite the PISA 2009 results the American students performed around the average (obtaining 502 points in science) according to OECD (2010b), research indicated that students had low scores in science interest (Do-Yong, 2006).

c. Factors associated with learning outcomes. In Europe, students' science achievement and student background have shown a strong relationship (Euridyce Network, 2011). Regarding interest in science, students with more advantaged socio-economic backgrounds or those who had a parent in a science-related career were more likely to have greater interest in science, and they were more willing to invest the effort needed to do correctly a science task in the test (OECD, 2007). The results of students from the United States are affected by their race, colour and social background (OECD, 2010a). The lack of motivation towards science can be explained in part because of teachers' over-dependence on textbook use, considering that over 90% of all science teachers reported that they rely almost entirely on textbooks for their classroom teaching (Do-Yong, 2006).

2.1.2. Australia

a. Structure. In most of the Australian state curriculum frameworks, students are expected to meet the standards over a two-year period, in many primary schools the first 10 grades have composite classes (such as year 1/2, 3/4), and the curriculum focus is in general science content and scientific skills development. From grades 10 to 12 it is divided into biology, chemistry, physics, Earth and environmental science (ACARA, 2011).

b. Student achievement in science. Using the measurement of PISA 2009, Australia has demonstrated a performance in science above the average, reaching 527 points (OECD, 2010b). Nonetheless, Tytler (2007) mentioned there was an increasingly negative student attitude towards science and a decreasing participation in post-compulsory science subjects especially in physics and chemistry.

c. Factors associated with learning outcomes. A study conducted by Lokan, Hollingsworth, and Hackling (2006) indicated the centrality of inquiry-based learning in Australian science teaching is related with students' positive outcomes. However, analysing video lessons they found in their study limited scope for students to formulate their own research questions, devise experimental procedures and analyse data because the practical work was largely teacher-directed.

2.1.3. Chile

a. Structure. Science education is delivered as general science (called "Natural Sciences") in the first eight grades of schooling -which would be equivalent to primary and secondary school in other countries, but in Chile it is known as primary school-, and then science is divided into the three main areas: biology, chemistry and physics in the four grades of high school. Some exceptions are technical and art schools that can finish science education at the second year of high school (Gobierno de Chile, 2009a).

b. Student achievement in science. Chilean student science achievement is are not very positive (Navarro & Förster, 2012). In international measurements like PISA, nearly one third of Chilean students were not able to handle basic scientific concepts, interpret verbatim information or obtain simple conclusions in familiar situations, while on average OECD countries have only 18% of their students in this initial level (Gobierno de Chile, 2007). The results in 2009 remained almost stable. Chile obtained a score of 447, 53 points below the average of the OECD countries which was 500 (Gobierno de Chile, 2010b) .

The national measurement of learning outcomes in this country, Sistema de Medición de la Calidad de la Educación (Quality of Education Measurement System, SIMCE) defines three similar achievement levels than PISA: initial, intermediate and advanced. The results in science indicated that in 2009 around 40% of the students were in the initial level, which has slowly decreased over time. In 2011 there were 35% of pupils in the initial level, 33% in the intermediate and 32% in advanced (Gobierno de Chile, 2012d). These results coincide with the ones from international measurements.

c. Factors associated with learning outcomes. In PISA 2006 and 2009 male students performed better in science than female students (Navarro & Förster, 2012). The gender differential is an important point to be taken into account because it is not the global tendency in the participant countries in the test. Actually, there are several countries where girls obtained better results than boys in the science area (Gobierno de Chile, 2010b). Nonetheless, in the Chilean national measurement the difference between male and female students in science is not statistically significant (Gobierno de Chile, 2012d). The results in international and national science tests are highly dependent on the income level of the students' family (C. González, Martínez, Martínez, Cuevas, & Muñoz, 2009; Navarro & Förster, 2012; Valencia & Taut, 2011). This is one of the most worrying aspects for the government, because although in general terms the country seems to be better than other developing countries, it is one of the least equally distributed in terms of learning and income level (OECD, 2010a).

Finally, teacher's quality has been proved to affect students' learning in the Chilean context. A direct relation between teachers' performance level of assessment result and students' learning outcomes was found using data from the national measurement (SIMCE) applied in 2008 (Alvarado, Cabezas, Falck, & Ortega, 2012). Similarly, Bravo, Falck, González, Manzi, and Peirano (2009) stated that high-evaluated teachers have a greater probability of having class groups with higher scores in the standardized national test.

2.2. Promoting science education: policies and strategies

In an attempt to promote science education, different countries have developed policies or strategies. The most relevant policies are curricular modifications and standards creation. The main strategy is focusing on science teachers' knowledge and skills. These are described below.

2.2.1. Changes in the curriculum

Curriculum innovations in science during the sixties and seventies in the United Kingdom, other European countries and the United States had little impact on the practices of science teachers (Welch, 1979). In Europe the outcomes of these innovations are, as yet, unclear (Osborn & Dillon, 2008). Nowadays the diagnosis is that both content and pedagogy in the current science curriculum are failing to engage young people with the further study of science. Then, it is desirable a modification of curricula's orientation not only to knowledge of science but also to knowledge of how science works. According to Osborn and Dillon (2008) the presence of both elements should be considered a must in any school science curriculum.

In European countries, the structure of the science curriculum varies in organisation and specificity, according to the country culture. For instance, although biology, chemistry and physics are distinguished as the main areas of teaching science, in Spain the curriculum is divided into 10 units for each of the science subjects, whereas in England there are four units for science as a whole, not divided the three mentioned areas. In spite of the differences, commencing by introducing basic concepts that are then revisited in depth in later years is a common point in all curriculums (but England). Changes in science curriculum across Europe were necessary because knowledge was usually presented in fragmented concepts, not interesting to students, especially to girls. Also, considering that a growing body of research has shown that most students develop their interest and attitudes towards school science before the age of 14, much greater effort has been put now in assuring an attractive science curriculum for primary students. Thus, the major curriculum shift has been towards inquiry-based approach to teach -and learn- science (Osborn & Dillon, 2008).

Although there is yet still no consensus about what constitutes inquiry (Barrow, 2006), in science teaching it is the most widely recommended approach to teaching. In the present research the inquiry-based teaching approach is understood as teaching science focused on students' process and results of creating science by them. It encourages students to use critical and logical thinking considering alternative explanations, through making observations and proposing questions and answers, checking resources such as books or scientific reviews in order to plan research, collecting, analysing, interpreting data and contrasting it with the previously found evidence, predicting results and explaining the supposed ideas and final results (Furman, 2008). In general terms and in different countries effectiveness of inquiry-based teaching has been demonstrated (Akerson & Hanuscin, 2007; Lebak & Tinsley, 2010) at least in students' science interest, attainment levels and teachers' motivation (Osborn & Dillon, 2008). From the teachers' perspective inquiry-based teaching is an effective approach because it provides opportunities to children to use and develop a wider range of skills such as working in groups, explaining their written and oral expression and having more open-ended problem solving experiences (Osborn & Dillon, 2008).

In Latin American countries and in Australia scientific literacy has been defined as the main purpose of science teaching (Tytler, 2007; UNESCO-OREALC, 2005), and it has led to changes in science school curricula (Bencze & Bowen, 2009). The most accepted and commonly used definition of scientific literacy according to Navarro and Förster (2012) is the one given by OECD (2009): the capacity of individuals to use scientific knowledge to identify questions, acquire new knowledge and be able to explain scientific phenomena, understanding the characteristics of science being involved in science issues as reflective citizens.

Specifically in Chile, the need for inquiry-based teaching has been established as a requirement for science education, not only in high but also in primary schools (Gobierno de Chile, 2011b). Recently a curricular adjustment in science education was approved by the government for primary science teaching which comprises from 1 to 8 grades. This modification was in terms of structure and content organization. Now, grades 7 and 8 are expected to be taught by biology, physics or chemistry teachers like in high school, not by primary teachers as it has been so far. For primary science four main units across all grades are now defined: sciences of life (including human body and health), physics, chemistry and sciences of the Earth and Universe. The scientific inquiry skills presented in the new curriculum are: observing and questioning, experimenting, planning and conducting a research, analysing and communicating evidence (Gobierno de Chile, 2012a). The high school science curriculum has not changed from 2009. It is divided into biology, chemistry and physics. Apart from the changes in the curriculum, another important modification affecting science education in Chile is the introduction of standards for all teachers including science teachers. The standards for teachers in other countries and also in Chile are presented in the next section.

2.2.2. Standards for science teachers

In the promotion of science education different countries have developed standards for teachers, convinced about the idea that no science innovation will be sustained unless systematic and on-going professional development is assured in science teachers (Osborn & Dillon, 2008; UNESCO-OREALC, 2010). For example, in 1995 the United States introduced the National Science Education Standards, which described guidelines for all grades of science teaching including planning, how teachers facilitate learning, assessment, promotion of a good classroom environment and their role in the school community (National Research Council, 1996). Another set of standards for teacher preparation was presented in 2003, on which all the institutions should base their decisions and all the student teachers must demonstrate their competencies consistent. A new version has been recently created by the National Science Teachers Association (NSTA, 2012).

In Australia there is a voluntary certification based on the National Professional Standards for Highly Accomplished Teachers of Science, developed by the Australian Science Teachers Association (ASTA, 2002, 2009). It includes professional knowledge, practice and leadership areas. They are oriented to teacher professional development and teacher recognition.

In Scotland, there is a set of standards for full registration in the General Teaching Council (GTE, 2006) which is a requisite for teaching in state schools. Even though they are not specific for science teachers, such teachers need to perform according to them. The standards provide a concise description of the professional qualities and capabilities that teachers are expected to develop in the course of induction to register and also a professional baseline which apply to teachers throughout their careers. Likewise, here there is a set of standards establishing benchmarks statements required for all programmes of initial teacher education created by the Quality Assurance Agency for Higher Education (QAA, 2000).

In Chile the scenario is becoming similar to the Scottish system. There is a general teaching framework containing standards which is the base for the National Teacher Evaluation System (Gobierno de Chile, 2003), that is compulsory for teachers in public schools to improve the quality of education (Bitar, 2011; Docentemas, 2010). Further, general standards were created for teacher education programmes to assess the knowledge of student teachers before graduation and, evaluate the quality of their professional performance when they were in-service teachers. The twenty one standards were organised in four areas: preparation, establishing a good classroom environment, teaching and teacher professional work in the institution and outside it, suggestions to rank student teacher performance in practical work using a three level rubric, and guidelines to collect the information (Gobierno de Chile, 2001). These standards were based in previous works developed in the United States by Dwyer (1994), Darling-Hammond, Wise, and Klein (1995) and Danielson (1996). Nonetheless, an important problem was not considered in their implementation: how to evaluate the standards fulfilment. Each teacher education institution needed to determine their assessment instruments and benchmarks. It was a complex process for which most of the institutions were not prepared for (Avalos, 2003).

Nowadays, due to their debatable impact, these standards have been specified for each of the three levels of education: pre-school, primary school and high school. It is expected that institutions which prepare teachers orient the design and evaluation of their programmes based on the standards (Gobierno de Chile, 2011b, 2012b, 2012c), though this is not compulsory and any evaluation of the implementation has been presented yet.

The standards for pre-school student teachers presented basic general skills, pedagogical and subject matter standards in the areas: visual and musical arts, language, mathematics, natural sciences, social sciences and also the development of autonomy, identity and partnership (Gobierno de Chile, 2012b). The standards in the science area for primary school stressed pre-service teachers' knowledge about students' learning of scientific knowledge and thinking skills, structure and function of live organisms, movement, matter and its transformations, Earth and Universe (Gobierno de Chile, 2011b). However, in this set of standards indicators for teacher's skills to teach science seem to be absent.

Finally, the standards for high school pre-service teachers present basic professional skills are presented firstly, followed by six areas of knowledge standards: language, mathematics, history and social sciences, physics, chemistry and biology (Gobierno de Chile, 2012c). In sum, there are references to pedagogy and content in these guidelines, but no orientations to the development of how to teach the contents. The knowledge that combine content and pedagogy has been named pedagogical content knowledge by Shulman (1986) and its development in science teachers is described in the following section.

2.2.3. Development of Pedagogical Content Knowledge

In terms of teachers' professional knowledge, three main types have been identified in literature review: Pedagogical Knowledge (PK), Content Knowledge (CK) and Pedagogical Content Knowledge (PCK). PK refers to teachers' knowledge about teaching practice that may enhance learning. CK is the teachers' knowledge about the topic being taught (Sevian & Gonsalves, 2008). The term PCK was introduced by Shulman (1986), referring to the special amalgam of content knowledge transformed by the teacher into a form that makes it understandable, hence its importance to teach well any subject (Appleton, 2006). For Shulman (1986), PCK includes analogies, illustrations, examples, explanations and demonstrations to reformulate the subject knowledge and make it understandable to the students. This is especially important in subjects such as science which are considered difficult for pupils (Davis, Petish, & Smithey, 2006). Although PCK conceptualization has been in debate (van Driel, Verloop, & De Vos, 1998), it has been described as highly adaptive, connected, fruitful, innovative knowledge and it is used to solve classroom learning problems (Treagust & Harrison, 1999). PCK has been a major field of study for over two decades (Loughran et al., 2007).

The importance of PCK in science teachers has been stated by Smith (2000) because explaining science concepts is extremely challenging for teachers. In several cases these concepts are inaccessible for students, and "transitional concepts" need to be developed and addressed to facilitate their understanding (Ogborn, Kress, Martins, & MGillicuddy, 1996). The creation of those transitional concepts needs PCK, because it is in the origin of devices to make knowledge more understandable (Shulman, 1986, 1987). Indeed, Treagust and Harrison (1999, p. 40) asserted that "without a repertoire of pedagogical content knowledge to recognize how the content can be explained appropriately to less informed people, teachers will be less equipped to do their work effectively".

Even though good levels of PCK have been reported in expert teachers (Shulman, 1986, 1987), there is concern about its level in primary or elementary science teachers because of their usual lack of CK as well (Appleton, 2006; Davis et al., 2006; Ginns & Watters, 1999; James & Scharmann, 2007; Rice & Roychoudhury, 2003; Treagust & Harrison, 1999; Trumper, 2003; van Driel & Abell, 2010). Actually, primary science teachers are recognised for not being science oriented (Appleton & Kindt, 2002; Cobern & Loving, 2002), avoiding science teaching when possible (Appleton, 2006; Euridyce Network, 2011; Tytler, Osborne, Williams, Tytler, & Cripps, 2008) and feeling uncomfortable and unqualified to teach science (Ginns & Watters, 1999). This has been demonstrated for instance, in Rice's study (2003) where the majority of teachers felt their subject matter knowledge was weak. This is a problem because teachers who feel insecure about science teaching tend to have a detachment from science and they might perceive teaching it as simply fulfilling an obligation (Cobern & Loving, 2002).

Science PCK acquisition has been studied mostly in secondary science teachers, but their preparation as teachers is generally quite different from primary school teachers (Davis et al., 2006). The former tend to be specialists in a science area (e.g. biology, chemistry, or physics), so their PCK basis is clearly defined by their CK. On the contrary, in places such as Australia, the United Kingdom and in many regions of South America such as Chile, primary school teachers are generalist, they have to teach in average eight subjects including science (Hume, 2012; Vergara & Cofré, 2008). It implies they must have a workable store of CK and PCK for each subject they teach (Appleton, 2006) and consequently be able to perform multiple tasks at the same time (Ferguson, 2008).

Therefore, many elementary school teachers use teaching strategies that are more adequate for other subjects than for science (Appleton, 2006). This is especially relevant to take into account in the context of the new requirements of elementary science teaching methodologies. For instance, inquiry-based teaching is a new requirement in primary science in Chile, while it was introduced in the United States more than twenty years before (Gobierno de Chile, 2009b; Rutherford & Ahlgren, 1990). Though in many countries this approach has been adopted ten years ago or more, in Chile it is still new for teachers in schools and it is even more challenging for teacher educators to teach how to teach based on it (Gobierno de Chile, 2005).

Likewise, lack of PCK in beginning and pre-service teachers has been described (Treagust & Harrison, 1999; van Driel et al., 1998). According to Onslow, Beynon, and Geddis (1992) pre-service teachers' acquisition of PCK at the university may not be sufficiently robust to apply it into practice. van Driel et al. (1998) suggested that PCK is developed through an integrative process rooted in classroom practice, implying that pre-service teachers or beginning teachers usually have little or no PCK when they finish their initial teacher education. Indeed, in their words "teacher training programs usually do not exert a major influence on science teachers' PCK" (p. 682).

Regarding the development of PCK in pre-service teachers, there are two contradictory views. On one hand, from the point of view of Shulman (1986), teachers cannot craft PCK and explanations until they are content experts and also expert pedagogues, which happens when they have several years teaching the subject. Also van Driel et al. (1998) mentioned the need of thorough and coherent CK and teaching practice for the development of PCK. Then CK has been positioned as strongly influencing teaching practice (Appleton & Kindt, 2002; Arzi & White, 2008; Carlsen, 1991, 1993; Faye, 2009; Geddis, 1993; Zembal-Saul, Krajcik, & Blumenfeld, 2002). For instance, Carlsen (1993) indicated in his study that when teachers taught topics in which they had greater content knowledge, they asked more demanding questions to students and gave them more opportunities to speak in the classroom. Likewise, in Appleton and Kindt's study (2002) teachers with stronger subject matter knowledge tended to employ more innovative teaching strategies. However, it is important that "knowledge teachers use in teaching does not start in university nor does end with graduation. Part of it can be traced back to their school learning as students and part is influenced by life out of school… it does not grow linearly over time" (Arzi & White, 2008, p. 245). Then, it is possible to assume that while science content knowledge is a necessary pre-requisite for effective science teaching, there is not a simple correlation between the science content knowledge of a teacher and their ability to teach that knowledge in school (Hume, 2012; Lloyd et al., 1998).

On the other hand, works from several researchers have proved that PCK is possible to develop during teacher education (De Jong, Van Driel, & Verloop, 2005; Gess-Newsome &

Lederman, 1990, 1993; Magnusson, Krajcik, & Borko, 2002; Stofflett, 1994; van Driel, Jong, & Verloop, 2002). This happens when teaching experiences are offered to student teachers (Bryan & Abell, 1999; van Driel et al., 1998). According to Bryan and Abell (1999), teachers' experiences shape how teachers see their practice and what they hear from their own practice, giving an interpretive viewpoint.

It has been stated that teaching experiences trigger PCK development because it helps in identifying usual problems in practice, in approaching those problems, creating solutions and making sense of the outcomes of teachers' actions (Bryan & Abell, 1999). Also, teaching experiences contribute to creation of teachers' beliefs about learning and teaching (Pajares, 1992). Nevertheless, student teachers or beginning teachers' teaching experience is not always altered by their beliefs about teaching (McDiarmid, 1990; Munby & Russel, 1992).

In sum, the studies in PCK development in teacher education have stressed the relevance of CK or teaching experiences which might lead in changes in teachers' knowledge and thinking. Considering the points of view described above, Duit and Treagust (2003) established the need for closing the gap between theory and practice in teacher education programmes to ensure pre-service teachers acquire PCK and are able to put it into practice when they teach subject matters such as science.

In this research it has been considered that modifying the practice through reflecting on beliefs or conceptions is possible when it is carried out as a thought-intensive task. Following Bryan and Abell's finding (1999), confronting and critiquing pre-service teachers' own teaching must be done carefully during the initial teacher education because they usually do not have self-confidence to reflect thoughtfully. This and other themes related to teacher education trends and strategies are described in the next section.

2.3. Trends in teacher education

Being adequately prepared in terms of course work in initial teacher training seems to be internationally recognised as critical for becoming an effective teacher (OECD, 2005; Shulman, 2006; Vaillant, 2009), and specifically an effective science teacher (Davis et al., 2006). Smith (2008) mentioned that science teacher education programmes are designed differently according to their views about teacher education. As in many countries preparing teachers is seen now with more challenges than before (Borman et al., 2009; Graber, 1996), in this section the views, structure and challenges for teacher education are presented as follows.

2.3.1. Views in teacher education programmes

Programmes of teacher training or teacher education in the '90s were based on the view that learning to teach was a process of acquiring knowledge about teaching and subject matter knowledge (Borman et al., 2009; C. Carter, 1990), but the integration of theory and practice would be a students' task based on their own efforts (Wideen, Mayer-Smith, & Moon, 1998). This model followed the idea that learning how to teach was an additive process that largely bypassed person and setting, giving no space for thoughts on the role of beliefs, theories or preconceptions in teacher learning (Feiman-Nemser & Buchmann, 1989). The assumptions of this called **traditional** or **old teacher education view.** Here, knowledge from course readings and lectures can be transmitted directly into practice, and pre-service teachers develop professional knowledge before experience rather than in conjunction with experience (Bryan & Abell, 1999; Russel & Munby, 1991). In this view, prospective teachers were expected to take a significant amount of content courses in the subject areas they will teach, which would allow them understanding how knowledge is constructed (Borman et al., 2009). This view was criticized, arguing that its fragmented character did not enhance student teachers to integrate their experiences in ways that would help them learning how to teach (Zeichner & Gore, 1990).

At the present time, learning to teach is mostly seen as a deeply personal activity in which the student teacher has to deal with their prior beliefs, the university culture, the school, the society and the demands from the teaching context (Wideen et al., 1998). Thus, the most important programme features helping beginning teachers to learn how to teach are:

constant support, working with peers and a systematic long-term message that provides direction for personal development (Graber, 1996). This view is supported by **constructivist theory** of learning (Vygotsky, 1986), which has been applied in teacher education giving relevance to teachers' thoughts and experiences for the construction of their knowledge (Fenshamp et al., 1994). Both views underpin the current structure of teacher education programmes.

2.3.2. Structure of teacher education programmes

Nowadays in Chile and other countries two different teacher education programme structure can be identified: the concurrent and the consecutive. The concurrent (also called integrated) is the most known structure, where teachers have an integrated curriculum with pedagogy, methodology and subject matter courses (Borman et al., 2009; Gobierno de Chile, 2005). It could be considered as approaching in essence to constructivist view, even if in its implementation many variations exist. The consecutive structure (or alternative) appeared in the early eighties to meet the need of teachers in difficult-to-recruit areas such as science. The student teachers often had previously taken four years in degrees related to the subject matter, and they were put in a "fast track" to obtain the teaching certification (Borman et al., 2009). This is closer to the old or traditional view of teaching.

A common point between these two structures is the convergence of theoretical courses and practical teaching experiences, although they vary in length, the career phase where they are presented and the linkage with skilled teachers or mentors (Borman et al., 2009). Most of the programmes in European countries have a significant practical component combined with other sources of conceptual learning (Feiman-Nemser, 2008). This dual form of teaching method has increased in professionally oriented trajectories because teachers need to apply theoretical concepts to the classroom (Borman et al., 2009; Hagger, Burn, Mutton, & Brindley, 2008; Vermunt & Endedijk, 2011). Contrarily, in Latin America usually the practical component is the last in time and duration (Vaillant, 2009), although in the Chilean system it has been slightly increasing since the nineties (Avalos, 2003).

2.3.3. Challenges and strategies in science teacher education

Learning how to teach in teacher education programmes is considered a complex process (Avalos, 2011; Geddis, 1993; Graber, 1996) and specially challenging in science teachers (Hume, 2012).

First of all, there are challenges that student teachers are bringing into initial science teacher education. Pre-service teachers often have limited ideas about what the process of teaching science will be like (Geddis, 1993) and about what to do instructionally with pupils' knowledge (Tabachnick & Zeichner, 1999; Zembal-Saul, Blumenfeld, & Krajcik, 2000).

Furthermore, science teacher education requires cognitive and emotional involvement of teachers individual and collectively, their capacity and willingness to examine where each one stands in terms of convictions or beliefs, their perusal and enactment of appropriate alternatives for improvement (Vermunt & Endedijk, 2011). Likewise, it is expected that teacher education programmes develop student teachers' skills to critique, adapt and design materials (Duncan, Pilitsis, & Piegaro, 2010) and also to argument, due to the need to emulate and facilitate argumentation in their future pupils (Euridyce Network, 2011). Besides, teacher education needs to prepare future teachers to deal with the diversity in pupils' backgrounds and the constant change of science knowledge. Teaching how to teach is an uncertain scenario considering that perspectives on good teaching and good education are constantly shifting (Vermunt & Endedijk, 2011).

In the context of Chile, there is an extraordinary pressure in science teacher education to create a better science teacher career and, in consequence, a better school science education (Vergara & Cofré, 2008). Although it is known primary science is key to developing attitudes towards science and curiosity in children (Osborn & Dillon, 2008; Tymms, Bolden, & Merrell, 2008), most of the efforts, material and human resources have been concentrated in high school science (Vergara & Cofré, 2008). As an example, in primary science teacher education, seven of the 44 institutions only have more than five workshops or courses related to science or science teaching. Indeed, in most of the programmes less than 7% of the study plan is related with science or science teaching (Vergara & Cofré, 2008). In high school teaching programmes this is usually between 30-60% of the study plan (Cofré et al., 2010). This is relevant because as mentioned in section 2.2.1, today teaching science through inquiry is a requirement in science education in this country. According to Davis et al. (2006), in order to teach science based on inquiry, teachers must hold a “strong understanding of and abilities with regard to science inquiry” (p.615). Yet, in the 35 programmes that entitle student teachers as science teachers in Chile, less than 6% of the curriculum is related to science inquiry activities in average (Cofré et al., 2010), which is very low compared to science teacher education in developed countries (C. González et al., 2009).

Taking into account that “teacher education programmes would do well to devote more attention to the manner in which teaching implies the transformation of subject-matter” (Geddis, 1993, p. 682), it seems reasonable to expect the impulse of appropriate and innovative models from teacher education institutions (C. González et al., 2009). In this sense, Clarke and Hollingsworth (2002) emphasised that teachers need a model based on reflection and enactment in these domains: their practice, the outcomes of their practice, their knowledge, attitudes and beliefs, professional experimentation and the sources of teacher information, stimulus and support. To learn better in initial teacher education Shulman (1987) gave a valuable lead, indicating that effective teaching is a skilled and purposeful activity involving pedagogical reasoning and action processes. These have been part of the teacher education programmes strategies until nowadays. For instance, reasoning was investigated by Eshach (2006) based on an inquiry events workshop, where teachers were encouraged to advance scientific reasoning processes and to draw generalizations. As a result, the participants understood that teaching children through this methodology made them better thinkers, remarking the relevance of acting and reflecting.

In another study the effectiveness of learning activities during teacher education were explored from the perspective of the student teachers. Hagger et al. (2008) concluded the participants considered as a real source of learning the study of experience of a lesson they taught and another teacher observed. Secondly, they learned from feedback given by a mentor or regular class teacher, followed by the observation of their own lesson taught receiving advices from the observers about planning the lesson. Further, these teachers mentioned conscious review of the student teachers’ practice focused on a specific issue or concern, and then from observing experienced teachers’ lessons. The least frequently sources of learning mentioned were university input and ideas derived from reading research or professional literature. From the point of view of researchers and student teachers the strategies that connect educational theory and practice are described as crucial in initial teacher education. Some activities with this objective are presented in the next paragraphs.

2.3.4. Connecting theory and practice in teacher education

As it was mentioned, in traditional teacher education the first experiences in real schools give teachers the settings to apply the skills, then, the institutions completed their mission providing the theory of teaching and CK (Wideen et al., 1998). This might imply institutions do not appear as responsible in helping students in this connection and transfer. This constitutes a problem according to Feinstein (2010) because although students' correct transference of general principles to specific circumstances -different from those in which they learned the principles- is desirable, it is a complicated and unlikely expectation from the educational psychology perspective (Schwartz, Bransford, & Sears, 2005). This is because practical knowledge usually remains implicit (Vermunt & Endedijk, 2011) and pre-service teachers need to make it explicit to analyse and apply it (Vergara & Cofré, 2008). Nonetheless, there are strategies such as early practices, using videos or microteaching that can help student teachers bridging theory and practice.

a. Early teaching practices

Lack of opportunity to be inmerse in science teaching in schools is a problem for pre-service teachers, due to their need to observe effective and credible teachers to model their own practices. In addition, pre-service and beginning science teachers need to apply their knowledge in practical situations with pupils, to establish connections between science education theory and practical field experience (Ginns & Watters, 1999).

Early teaching practices are recommended to strengthen pre-service skills and explore the way of making sense of the theory in the practice (Mellado, 2003). According to Bryan and Abell (1999), offering early teaching experiences to pre-service science teachers in order to create meaningful opportunities for reflection and facilitate their learning from this experience is needed. In those experiences, the student teachers initially observe the actions of an expert teacher in a classroom, and over the time they are scaffolded into increasingly more central professional classroom teaching opportunities provided by the expert (Hume, 2012).

In Bryan and Abell's research (1999) one case study pre-service science teacher reflection process was followed to understand how her experiences within the context of reflective science teacher education influence the development of professional knowledge. The researchers made the participant analyse her own and others' practice, compare her actions in the classroom to her vision of teaching science and confront tensions in her thinking about science teaching and learning to resolve these tensions. As a result, the student teacher became aware of her beliefs about teaching and learning from experience and accompanying reflection on the teaching experiences. At the same time, the explanation of her beliefs about science teaching provided a reference point for analysing her practice. Although these findings are descriptive-interpretative only based on the study methodology, there are correspondent with other studies where reflection stimulated reframing and revising practice (Harford & MacRuairc, 2008) or provided professional learning to new teachers (Hagger et al., 2008).

b. Video usage in teacher education

Using video cases in teacher education has been recognised as an important tool for several years (Dawson, 1975; Kallenbach & Gall, 1969; Spelman & John-Brooks, 1972). Now, there is agreement that it allows the observation and evaluation of teaching situations, supporting student teachers because of the opportunity to discuss teacher performance or teaching styles before being teachers (R. J. Beck, King, & Marshal, 2002; H. Kim & Hannafin, 2008; Kurz & Batarelo, 2010; Moreno & Ortegano-Layne, 2008; Sonmez & Can, 2010).

For instance, Dawson (1975) in a study with two students teachers indicated that videotape feedback of teacher classroom performances tended to change specific behaviours, while in the control teacher they remained stable over the time. The author pointed out the utility of using videos despite its expensiveness, especially when other forms of supervisions had failed. However, due to the small sample size of this study it is difficult to assert the findings were consequence of using videos, focused supervision or by chance.

Different approaches can be used while implementing videos with student teachers (Sonmez & Can, 2010). For instance: leading a whole class open-ended analysis guided by student teachers observations after viewing the video and progressively guiding the reflection (Kurz & Batarelo, 2010). Alternatively, using short fragments to emphasise particular teaching aspects or illustrating good practices (Star & Strickland, 2008). Videos can be used also to allow pre-service teachers watching successful and failed teaching practices, in order to learn from other teachers' good decisions and mistakes (Darling-Hammond & Hammerness, 2002).

Video analysis has been positioned a powerful tool to promote analytical reflection in pre-service teachers (Bencze, Hewitt, & Pedretti, 2001; Boling, 2007; Harford & MacRuairc, 2008; Kurz & Batarelo, 2010; Whitehead & Fitzgerald, 2007). Videos develop a shared language to discuss what the student teachers observed in the same grounds (Sonmez & Can, 2010) and because it avoids the impact of external face-to-face observers that may be felt negatively evaluative (Newhouse, Lane, & Brown, 2007). At the same time, video analysis promotes the culture of observation and critical dialogue in education (Harford & MacRuairc, 2008), and it empowers teachers to recognise and critically evaluate others' practice (Loughran & Russell, 2002). Comparing videos with teaching observation, the videos present more benefits than observing in vivo due to its real-time nature it does not allow student teachers to replay to deconstruct the practice (Harford & MacRuairc, 2008). The characteristics of videos that make them more cognitively salient and more effective as teaching-learning tools are: their authenticity, dynamic moving quality and visual explicitness of images. Besides, according to Beck, King and Marshal (2002), pre-service teachers' own videos might be more effective in bridging theory and practice, if they use previously learned theories and concepts to focus and interpret the videos.

In science teacher education success using videos has been demonstrated when it relies on the teachers' ability to recognise dynamics, strengths, weaknesses and teaching skills (Sonmez & Can, 2010). This ability comprises (a) identifying what is noteworthy in the teaching situation, (b) making connections between the specific classroom interaction and the present principles of teaching and learning, and (c) using what one knows to reason about classroom interactions (Sherin, 2005). Sonmez & Can's study (2010) investigated this skill in pre-service science teachers using videos. They showed progress through systematic assignments and discussions. These researchers warned there would be little benefit of using such videos in the absence of teacher observation and evaluation skills.

c. Microteaching in teacher education

Microteaching is a short duration teaching experience, often around 5-15 minutes (Kpanja, 2001; Mohan, 2007). It is a common practice in the United States science teacher education (Ferguson, 2008) and also in Nigeria (Kpanja, 2001). During microteaching, trainees take turns teaching a lesson in front of their peers, usually in four steps: (1) the briefing or orientation, (2) teaching the lesson, (3) the critique or discussion and (4) re-teaching the lesson (Orlova, 2009). Peers are usually playing the role of primary or secondary students, who should make questions and wrong interventions to facilitate the teachers' deployment (Pauline, 1993).

In theoretical terms, microteaching has been presented an efficient and effective technique in teacher training programmes because the simulated context might give a teaching experience to make the pre-service teachers aware of the various skills of which teaching is composed. They can focus their attention in clearly defined aspects of their teaching, removing the problem of control or discipline that would be distractive with real pupils. Here, video recording the microteaching episode, peers and tutor feedback to stimulate self-analysis has been recommended (Mohan, 2007). In the same argument, Ferguson (2008) asserted that observing, analysing and discussing classroom performance was enhanced by the use of videotaping microteaching episodes, and it could help student teachers to see themselves from a different perspective.

In the nineties the need for systematic empirical studies on the effects of microteaching in teachers' beliefs, attitudes and thinking was stated (Gess-Newsome & Lederman, 1990). In Chile, the relevance of microteaching in science teacher education has been highlighted, but it is still an uncommon practice in teacher education programmes (Vergara & Cofré, 2008). These researchers declared that performing science teaching and being recorded in videos would allow student teachers analysing their practice, re-constructing their theories and strategies. Also, it may generate a metacognitive reflective process in which student teachers could be aware of their conceptions, attitudes and how those are conducting their practices.

A little recent past empirical research in microteaching usage during pre-service science teacher education has been found. Some studies were reported several years ago, such as the one from Sparks and McCallon (1974), who investigated if pre-service primary science teachers enrolled in a microteaching laboratory experience would have more positive attitudes toward science teaching than another group taught with traditional science methods and a comparison group. The groups were already constituted. They recorded the experimental group performing a microteaching episode in a small group of children, and then the teachers viewed the video tape and criticized it with support from the professor or his assistant. This was done six times. To access student teachers' attitude change towards science teaching, they administrated questionnaires to 98 teachers in a pre-test and post-test eight weeks later, containing a semantic differential scale. They concluded there was a more positive attitude in the group that took a regular science methods course than in the group having a microteaching, although in both groups it was better after the microteaching than before performing it. In the control group the attitude towards science was more negative. The main problem in this study was the usage of qualitative ordinal data as a quantitative continuous variable. The researchers based their conclusions on means comparisons and standard deviations, which is questionable when using semantic differential as the unique source of data.

One of the only recent works has found benefits of using microteaching in pre-service science teachers was presented by McLaury (2011). He looked for teachers' perceptions of microteaching assignment performances in connection with their beliefs about teaching science. He found that teachers' beliefs, rather than instructor or peer-based assessments, served as the primary determinant by which they perceived personal success in microteaching. Also in this study, explicit instructor-planned challenges to teachers' beliefs were generally rejected as sources of change. Intra and interpersonal interactions apparently resulted in the creation of pedagogical content knowledge. However, many other interactions were devalued or ignored for a variety of personal, experientially based reasons, reaffirming the fundamental role of pre-existent belief systems in the selective creation and processing information of teachers' knowledge.

In other areas of teacher education benefits of microteaching have been reported, such as allowing student teachers to distance themselves from their teaching, helping them to notice and respond to both strong and weak aspects of their teaching and motivating them for teaching (Maclean & White, 2007; Orlova, 2009). Indeed, for Orlova (2009) video recording teachers' practices is considered one of the most valuable tools because it provides an objective and permanent source that can be viewed repeatedly to observe various aspects of classroom practice, which can lead to an improvement in self-awareness when teachers reflect on the episodes. Likewise, Ferguson (2008) has indicated microteaching can be useful to expand science pre-service teachers' notions of teaching expertise. In a study carried out by Yerrick, Ross, and Molebash (2005), pre-service science teachers were videoed in schools and asked to edit the videos focusing in children's beliefs identification. With this input teachers modified their lesson plan and then recorded the lesson for those pupils, editing again their videos focusing in their actions conducting to learn. They included in the video personal reflections about the experience that were thereafter analysed by the researchers in group discussions. Teachers' beliefs shifted regarding children's thinking and changes in teachers' planning and their instructions were reported. Though, the researchers stated because of the exploratory design they were not able to verify validity or reliability of their findings.

In the context of Nigeria, a research using video-recorded microteaching was carried out with 20 pre-service teachers in an experimental group and a comparison group with other 20 teachers who had non-recorded microteaching. Both groups discussed their performance with the lecturer in charge. The researcher concluded the group that used videoed microteaching showed more significant progress in their teaching skills and they behaved more confidently and positively towards microteaching than the control group (Kpanja, 2001). Although this research stressed the importance of microteaching and videos, the author neither reported which teachers' skills improved nor how they were measured. Also, there was an ethical issue involved here because student teachers were not told they were participating in a study. Then, it is not surprising he found members of the control group were less enthusiastic and feeling to be inadequately prepared for subsequent teaching.

Microteaching experiences have been also used for promoting reflection in pre-service teachers (l'Anson, Rodrigues, & Wilson, 2003). These researchers used microteaching in different teaching subjects such as religious, moral and philosophical studies, information and communications technology (ICT) and science education. The researchers considered the student teachers' analytical process involved in reflection as pre-critical, internalised and hypothetical reflection thresholds. Also, they stated in order for reflection to occur pre-service teachers needed to be encouraged to view their development from various angles, which they called “multiple refractions of student experience”. The triangulation of a microteaching episode by the teacher, peer, and tutor/teacher’s eyes was a significant aspect in scaffolding the development of the process of reflection. In this study the researchers did not specified if their findings were different according to the teachers’ subject or if they improved their teaching skills.

Another study was carried out in Australia by Ginns and Watters (1999). Three case studies were chose from a sixty one sample that completed a test of efficacy beliefs one year before the study. The researchers recorded student teachers teaching science in real contexts during several lessons, and coded their behaviours using criteria obtained from literature review. A questionnaire was administrated to investigate in their beliefs and also the test of efficacy beliefs. They used the mean of the whole school staff application of this test in one of the schools to compare student teachers’ progression, and concluded their case-study teachers were more homogenous because they showed a minor standard deviation than the whole group of the school. Although in the qualitative results they remarked the three teachers valued the peers’ feedback received in their efficacy, in terms of quantitative analysis it is questionable the sampling process and the comparison with a group not necessarily similar. The researchers indicated the participants were selected according to high and low score in the test but also considering the convenience criteria from researchers’ home distance to the school. This might imply bias in the results obtained because it seems there was not random selection or measurement of other variables in the teachers selected, and they did not compare the results with other teachers of the same group who were not recorded. Moreover, these researchers did not mention what criteria from the literature review they chose to assess teachers’ practice or how they did it.

Despite previous findings, long time ago a study conducted by Kallenbach and Gall (1969) found microteaching had the same results as more traditional teaching methodologies. Similarly, Copeland (1975) did not find statistically significant differences in teaching skills between teachers’ simulated microteaching episodes and classroom experiences.

Although according to Pauline (1993) the main critique of the microteaching setting is its artificialness (it would not be sufficiently comparable to the classroom for transfer of skills), all previous work recommended it as a valuable teacher education technique to prepare teaching skills, or even as the most effective (Kpanja, 2001). Equally, in some of the studies where microteaching was not found to result in significantly higher ratings of teacher effectiveness, it was remarked as an efficient training strategy since it achieved similar results when compared with other strategies but only in one fifth of the time and with fewer administrative problems (Kallenbach & Gall, 1969). Thus, efficiency of microteaching seems to be an important point to consider especially focused on the transferability of the skills developed or knowledge gained into a different context. Any of the studies found in this literature review proved what skills can be generalised and which others cannot or the factors affecting the transference into real teaching contexts.

A common point between these studies to increase the power of microteaching is that peers and or tutor feedback have been recommended to stimulate teachers' self-analysis (l'Anson et al., 2003; Mohan, 2007). Peers' feedback and assessment in teacher education are described in depth in the next section.

2.4. Peer Learning and Assessment in teacher education

In most of the studies cited in the previous section, student teachers' microteaching was accompanied by feedback from and to peers, which is based on peer learning. The definition of peer learning, peer assessment, the principles underpinning it and peer assessment critiques are presented in this section.

2.4.1. Definitions

Peer learning is understood as the use of teaching and learning strategies in which students learn with and from each other without the immediate intervention of a teacher (Boud, Cohen, & Sampson, 1999). A similar definition has been presented by Topping (2005), who indicated peer learning is the acquisition of knowledge and skills through active helping and supporting among status equal or matched companions. It often involves learners from similar social groupings helping each other to learn and learning themselves by doing it. In this sense, peer assessment could be considered as a peer learning strategy in which assessors might identify or create the assessment criteria (Boud et al., 1999). When the assessment is based on performance, criterion-referenced assessment (where different levels of proficiency for each criterion are defined) is appropriate to give feedback (Sluijsmans & Prins, 2006).

Topping and Ehly (1998) defined peer assessment as an arrangement between peers to consider the amount, level, value, worth, quality or successfulness of the products, outcomes or learning from other similar status learners. It can be quantitative, qualitative or both, summative or formative or both, face-to-face or remote (McLuckie & Topping, 2004). In peer assessment, feedback implies an evaluation of the practice being watched to construct the critique, corrections or suggestions to give to peers (Kpanja, 2001). According to Topping (2010) there are different types of feedback: directive and facilitative, norm-referenced or self-referenced, formative feedback (positive or negative in content), among others. It has been said in general, feedback that lacks specificity or is too directive could be damaging, while too long feedback perhaps could be ignored. Feedback from peers can be more immediate, available timely and personalised than teacher' feedback. In this research peer assessment was formative, qualitative and feedback was given face-to-face between student teachers. In the first moment of the PA intervention it was self-referenced and further it was norm-referenced.

2.4.2. Underpinning theories

Students might learn better from another student or students who have a similar understanding of the material to learn because they are around the zone of proximal development and trigger their peer's learning (Vygotsky, 1986), then, a process of scaffolding occurs between students with similar cognitive characteristics. Indeed, Nicol and Boyle (2003) mentioned that discussions between students in the same level of knowledge can enhance students' own mental knowledge model construction process through negotiation of meaning, which can help them to achieve an improved conceptualisation. Negotiation of meaning is a construct that could serve to explain the possible success of peer formative assessment in teacher education. In the perspective of Moje et al. (2001), negotiation of meaning needs the construction of a third space in between everyday knowledge and teaching knowledge, where language, literacy and science learning in diverse classrooms could be developed. This space provides the mediational context and tools necessary for cognitive development (Gutierrez, Baquedano-Lopez, Tejeda, & Rivera, 1999). Negotiation of meaning runs in the third space of understanding, where student teachers can jointly redefine elements of good teaching with their peers through assessment and discussion. It is constructed when four characteristics of interaction are present; (1) drawing from student's everyday discourses and knowledge, (2) developing student's awareness of those discourses and knowledge, (3) connecting them with the science discourse and negotiating the understanding of both discourses and knowledge, so that they not only inform the other, but also merge to construct a new kind of knowledge based on the negotiation of meaning (Moje et al., 2001). Here, as Catalán (2010) asserted, the construction of meaning is integrated from the construction with peers.

Another possible route for peer assessment in student teachers to work is because it would allow self-regulated learning, by giving them the opportunities of talking about their own decisions, beliefs and practices (Vermunt & Endedijk, 2011). This is aimed considering the importance of life-long learning and continuous professional development. Students should gradually become the owners of their learning process, and they learn better when they actively construct their own knowledge learning in interaction with fellow students

(Bakkenes, Vermunt, & Wubbels, 2010). However, self-regulation of learning has been seen as a necessary but not sufficient condition to develop pre-service teachers' beliefs and skills they need (Vermunt & Endedijk, 2011).

2.4.3. Empirical studies using Peer Assessment

Just a few studies focused on the role of PA in science teacher education were found. This is why other areas where it has had an application are discussed also in this section.

In science teacher education, a study conducted by Gess-Newsome and Lederman (1990) looked for pre-service teachers' perceptions of teaching, instructional decisions and changes in beliefs. They used microteaching in 17 student teachers in a case study design. They received informal feedback from peers and formal written feedback from the course instructor and also received their videotape to write a self-critique. Student teachers completed a questionnaire before the microteaching episodes and after them. In spite of all the sources of data this study had, the authors reported only concerns of student teachers at the moment of performing. They identified 12 areas of concern, grouped into concerns about themselves and about students, informing an apparent shift of focus from "concerns for self" to "concerns for students". The researchers did not go further in the effects of feedback on the student teachers' beliefs. Moreover, the study did not describe the role of PA in microteaching experiences. Thus, it seems the findings reported are inconclusive for the research objectives.

In Korea, a study with 82 pre-service teachers enrolled in an Educational Technology course investigated if having an structured assessee's role had an impact in metacognitive awareness of their learning process, their performance and their motivation towards PA (M. Kim, 2009). A metacognitive awareness questionnaire and a motivation survey were completed before the intervention. The performance was measured in an assignment to create a concept map on instructional design that all the participants submitted for peer feedback and tutor marks. After receiving feedback, the participants were randomly assigned to the experimental condition that received a back-feedback opportunity (n=40) and the control condition that did not have that opportunity (n=42). The back-feedback was to reflect and give their opinion on the peer feedback obtained. After the revising task, students teachers resubmitted their concept maps, completed the metacognitive awareness questionnaire and survey as the post-tests. Results showed that the experimental group had higher metacognitive awareness, performance and better attitudes towards PA than the control group. The last result can be explained because of the feeling of unfairness in the student teachers in the control group that observed the back-feedback activity in the others.

In Taiwan, a networked PA system was implemented to improve the quality of 24 pre-service science teachers' inquiry-based activities and also investigate the correlation between peers and experts' evaluation of those activities (Tsai, Lin, & Yuan, 2002). The evaluation was based on three criteria: creativity, relevance and feasibility. The researchers concluded pre-service teachers improved the quality of their science activity design as a result of PA based on an improvement in the score from peers' and experts' view. Nonetheless, the agreement between peers' and experts' marks was not high enough as their expectations. Here, the same problem than in Hume's study (2012) is noticed. The researchers attributed a causal relation between results and PA, but they did not have a comparison group to control other variables influencing the emergence of the results. Also, it is not surprising the correlations between peers' and experts' marks were not high, because prior discussion or agreement in the criteria used was not reported. Then, it is not possible to assure the validity of the assessment in this case. Clarification or negotiation of rating criteria is crucial in PA (Boud, 1990; Orsmond, Merry, & Reiling, 2000; Sluijsmans, Brand-Gruwel, van Merriënboer, & Martens, 2004).

Another study in a university in Turkey looked for PA of elementary science teaching skills in pre-service teachers, using microteaching episodes (Kilic & Cakan, 2007). The assessors and course instructor evaluated twice the episode through a PA questionnaire in a Likert Scale between very good and very poor (comprising science content and teaching knowledge, teaching-learning process, class management and communication). The researchers concluded student teachers can be reliable assessors because their scores were more correlated to the instructor score the second time, but the number of peer raters should be around five to sustain an acceptable reliability. They recommended PA to be conducted from the first year of teacher education to develop assessing skills. This study did not describe if assessed teachers improved their skills the second time they performed because it was not part of the study aims.

In other areas of teacher education the interest on PA is increasing (Woolhouse, 1999), because critical evaluation of peer performances is itself important in teacher training (Sluijsmans, Brand-Gruwel, & van Merriënboer, 2002). For instance, a study conducted in the Netherlands by Sluijsmans et al. (2004) aimed to develop the skill to define performance criteria in 93 pre-service teachers splited in half and randomly assigned to control and experimental groups. The experimental group was trained to develop the skill while the control group was not. The study used PA questionnaires, rating forms, student questionaires and interviews. Among the main findings, student teachers in the experimental group were more capable in using the set criteria determined during the PA training than teachers in the control group, they used the criteria more often and they felt more able to assess than before. However, the researchers did not find statistically significant improvement in their performance. The researchers concluded PA as a skill can be trained, and further follow-up studies are suggested in a very related skill: giving feedback.

Another research underlined the relevance of peer formative feedback. Harford & MacRuairc (2008) enganged 20 student teachers during their first teaching practices in peer videoing and analysing in pairs their teaching. As each student teacher was videoing and videoed, a greater understanding of and empathy with the tensions and challenges of the process was possible. Among the findings the authors indicated that student teachers developed their reflective skills and peer feedback had an impact on their classroom practice, bridging the gap between reflection and practice. The participants valued their exposure to a range of diverse teaching methodologies, the transference of teaching skills from one subject area to others, and the utility of seeing teaching methodologies in operation and not only in a list of recommended strategies. They made links between the practice viewed and their future work evaluating the project as a powerful mechanism for conducting self-review and dialogue regarding classroom practice. They appreciated informal and formative collaboration, remarking that structured assessment would reduce their engagement with the process and the quality of the reflective dialogue, which contradicts another study where structured PA was positively viewed by the participants (Orsmond, Merry, & Reiling, 1996).

PA has been also applied in in-service teachers. An example is the study conducted by Wen and Tsai (2008) with 37 science and mathematic teachers who submitted master's thesis proposals in a web-platform. The teachers in this study used the assessment criteria in several projects before assessing their peers' work which was also marked by an instructor. PA was made in three rounds. The researchers noticed an increase in peers and instructors' scores, even though they were medium or low correlated. The quality of the provided feedback improved and the student teachers showed more knowledge in research methods. However, in this study peers and instructor's scores were not statistically consistent in two of the three criterias evaluated, which might make the findings doubtable.

Bakkenes et al. (2010) conducted a year longitudinal study with 94 in-service secondary teachers in diverse areas, who reported six learning experiences using digital logs. Content analysis was applied to the experiences to identify and associate learning activities and outcomes. The authors stated that teachers who participated in reciprocal peer coaching and collaborative projects often reported they got new ideas from these, little experiences of negative emotions and qualitatively better learning outcomes than informal learning in the workplace. Even so, these findings need to be considered with caution because the associations were based on self-reports only.

Apart from teachers, benefits of peer learning in other students have long been recognised (Topping, 2005). To mention some of them; taking responsibility for students' own learning and deepening their understanding of specific course content (Boud et al., 1999) or in the development of linguistic competence and the self-concept as a writer in the case of school students (Duran & Monereo, 2008).

2.4.4. Critiques to Peer Assessment

Regarding the critiques of PA, Boud et al. (1999), mentioned the main one is its validity. Usually low or medium correlations between with peers assessors and experts' marks are found (Tsai et al., 2002; Wen & Tsai, 2008), even so in some studies medium-high correlations have been also identified (Kilic & Cakan, 2007). In this regard, Topping (2005) asserted this type of correlations between peers and experts' marks are referring to reliability of PA and not to validity of PA. Related with this issue, MacArthur, Schwartz, and Graham (1991) have indicated that discussion, negotiation and joint construction of assessment criteria between learners is likely to increase reliability and sense of ownership in the criteria used. In the same line of thought, Topping (2010) affirmed PA is more reliable when is supported by training, checklists, demonstrations, teacher assistance or monitoring.

Between the limitations, PA needs a carefully designed setting to be implemented, and sometimes it can be highly time consuming (Boud et al., 1999). Likewise, Lin, Liu, Chiu, and Yuan (2001) have observed that PA is effort consuming and extremely low or high scores given by assessors without adequate knowledge need to be taken into account because they might affect its validity. Also, peers might present problems to evaluate their friends' work, although it can be avoided in anonymous evaluation (Wen & Tsai, 2008). Another problem is the difficulty of using PA in personal reflection or interpersonal relationships between students, as it usually requires a conscious meta-cognitive development to recognise or verbalize the reflective processes at the moment it is occurring. Also, attention must be paid to inappropriate forms of assessment could motivate students to take a surface approach of learning instead of meaningful (Boud et al., 1999). These authors affirmed that structured assessment needs to be taken into account to solve this difficulty. Similarly, Topping (2010) highlighted PA heavily demands the communication skills of the assessor and assessed, and he also mentioned the emotional component of face-to-face feedback can affect reliability of the assessment.

In sum, most of the critiques have referred to PA based on quantitative marks, and the problems with some of the empirical studies are attributable to students' lack of experience or appropriation of the criteria to assess. Nonetheless, PA has been recommended as an effective but under-utilised type of formative assessment, as a tool to promote reflection (Topping, 2010). As reflective thinking is one of the expected skills to be developed in initial teacher education, its definition and role in the present research is discussed the following section.

2.5. Reflective thinking in teacher education

The first author who developed theory about the teacher thinking concept was John Dewey, whose ideas about systematic and persistent analysis were later expanded by Schön as cited in Orlova (2009). Schön (1991) stressed the relationship between reflection and experience, differentiating the reflection in action (during the practice) and the reflection on action (after the object of reflection was performed). Reflective teaching has been defined as the critical exploration of a teacher's own teaching practice and it applies to all educational situations (Wallace, 1998). Actually, a teacher who is a reflective practitioner continually evaluates the effects of their choices and actions, and this is a way to develop their professionalism (Davis et al., 2006), their knowledge and practice (Gunstone, Slattery, Baird, & Northfield, 1993). In fact, professional expertise can be developed in professional action (Orlova, 2009).

Nowadays, reflective teaching in teacher education encourages critical reflection during the whole teaching process, as a critical self-assessment of the teaching skills that student teachers require. These skills are necessary to thoughtfully analyse and determine how their own belief system and attitudes impact on their decisions and actions in the classroom. The critical reflection should allow teachers to develop the skills to analytically and objectively consider teaching processes as a way of improving classroom practices (Orlova, 2009). Engaging students in reflection is important for talking about learning and teaching, identifying inconsistencies between beliefs and practices and motivating them in inquiry-oriented science teaching (Arellano et al., 2001; Crawford, 1999; Sillman & Dana, 2001; Sweeney, Bula, & Cornett, 2001).

Empirical studies such as in Harford & MacRuairc (2008) have shown that although student teachers are often aware of the importance of appearing reflective, they usually do not see its application to the real life teaching experience. Furthermore, reflective practice appears as an individual, isolated action, and consequently, there are a few studies focused on the importance of colleagues or peers in the reflection (Davis et al., 2006).

As a result of scaffolding reflection, Harford & MacRuairc (2008) found in their study a clear impact on critical discussions between pre-service teachers, indicating "they identified the incremental nature of the level of critique expected as the study progressed as a significant contribution to the depth of analysis and their competence to engage critical discourse related to their practice” (p.1889). This was percieved by the student teachers as well. Nonetheless, those findings need to be taken considering the student teachers gave these comments in an open tutorial which was seen by the tutor, then, its presence may have impacted on the candidness of the student teachers’ views expressed. Also, because the evaluation of reflection as deeper, more critical and meaningful was done qualitively by the monitors of the programme only and not in comparison with the participants’ perspective.

Helping teachers to think about their own practice and colaborate with colleagues in this purpose was a practice with powerful effects according to Clark (1988). Teachers in his study reported that describing their teaching plans and intentions, explaining their reasons underpinning the actions and decisions in classrooms gave them a new meaning in their teaching. However, the researcher remarked this reflection usually required that teachers stop and think, finding words and reasons to justify their thougts and beliefs, take a second look at themselves and their teaching, a situation that is not very common during or after teaching practice.

It has been proved that when given a chance to reflect and confront pre-service teachers’ beliefs or conceptions, they can develop a deeper understanding of teaching and the teachers’ role (Bryan & Abell, 1999). In terms of setting to promote pre-service teacher exchanges through critical reflection, providing a safe environment to share their ideas among peers and with teacher educators seems to be a significant factor (Graber, 1996).

Through reflection on practice, teachers can also change their beliefs (Mansour, 2009). Indeed, in a study carried out by Gunstone et al. (1993) the extent of change in pre-service teachers’ beliefs and knowledge was dependent on how reflective they were. Recognising the importance of teachers’ beliefs or conceptions in teachers’ thinking, a detailed description is presented in the following pages.

2.5.1. Teachers' beliefs or conceptions

According to Eshach (2006), teachers' beliefs system is important to take into account in science teaching, because it may explain teachers' behaviour toward science teaching. Likewise, this system might influence the science teacher's thoughts and actions (Czerniak & Lumpe, 1996). The role that teachers' thinking play in their practice has been widely explored (Isikoglu, Basturk, & Karaca, 2009), especially in science education (Eshach, 2006).

Even so, the terms "belief" or "conception" have different connotations in educational research (Mellado, 1998; Pajares, 1992). In the present study, educational belief is understood as a set of representations guiding teachers' concept of learning-teaching and their own role on the process (Ruys, Van Keer, & Aelterman, 2010), while conceptions are focused on a specific topic in the teaching process (Hermans, van Braak, & Van Keer, 2008). The characteristics of student teachers' beliefs have been examined, usually using them to inform the effects of subsequent experiences in pre-service programmes (Wideen et al., 1998). Also, there is a wide body of studies involving teacher conceptions about learning, teaching and student learning (Vermunt & Endedijk, 2011). Other researchers have focused directly upon the characteristics of teacher prior beliefs or in teachers' self-efficacy beliefs (Appleton & Kindt, 2002; Ginns & Watters, 1999; Gunstone et al., 1993; King & Wiseman, 2001; Loughran, 1994; Scharmann & Orth Hampton, 1995; Wideen et al., 1998). Although this is the area with the largest body of research and self-efficacy beliefs have been associated with effective teachers (Scharmann & Orth Hampton, 1995), teacher efficacy is an elusive construct (Tschannen-Moran & Hoy, 2001). According to these researchers, "a teacher self-efficacy belief is a judgement of his or her capabilities to bring about desired outcomes of student engagement and learning" (p. 783). How to measure this kind of beliefs have been also in debate. Some have argued that measuring the teachers' general attribution of the pupils' result is an approach to self-efficacy (Ashton, 1984; Guskey, 1982) , and some others that self-efficacy are content and context specific beliefs (Gibson & Dembo, 1984). For instance, Bandura (1997) established that teacher sense of efficacy is not necessary uniform across the different teacher tasks they have to perform. He created a scale divided into decision making, instructional, disciplinary, community involving and the efficacy to create a good classroom climate, but information about its reliability has not been available (Tschannen-Moran & Hoy, 2001).

Specifically in science education, teacher efficacy beliefs have been measured traditionally with the Science Teaching Efficacy Belief Instrument (STEBI) (Enochs & Riggs, 1990). Bearing in mind that most of the previous work has been developed in English speaking contexts and the participants of the present research were Spanish speakers from Chile, the possibilities to find an instrument already tested and or adapted to this population were reduced. The work from S. Rodríguez, Núñez, Valle, Blas, and Rosario (2009) is one of the few available in Spanish. Nonetheless, this group of researchers did not develop an instrument but translated the Teachers' Sense of Efficacy Scale by Tschannen-Moran and Hoy (2001). As information about the adaptation process and the validity and reliability of this version was not presented -as there was for other countries by Klassen et al. (2009)- , it was not considered as a suitable instrument to apply in this research. Thus, in spite of the importance of self-efficacy beliefs in teachers' conceptions and practice, practically it was not possible to count with an already available valid measurement to apply in the research participants.

Nonetheless, in the present research the influence of beliefs in the construction of professional knowledge in student teachers was taken into account. Pajares (1992) and Richardson (1996) have developed reviews of research in the field of beliefs held by beginning teachers before entering training programmes. Their work showed a paradoxical issue; teachers' prior beliefs are often seen as problematic by researchers or teacher educators. This is because beliefs act as filters, and they might act as barriers to the student teachers' receptivity to the knowledge being offered in a teaching program (Wideen et al., 1998). However, as they are the origin of teachers' knowledge about learning and teaching, they are necessary to be transformed to create new knowledge (Pajares, 1992). This author stated that student teachers' beliefs play a crucial role in their interpretation of professional knowledge. They are usually established and fixed when student teachers were in college and changes during adulthood are uncommon.

Consequently, Richardson (1996) proved that conceptions about teaching come through personal experience, schooling, instruction and formal knowledge. This perspective is enriched by other studies, which proposed that the origin of teachers' beliefs in exposure to cultural archetypes of teaching or years of pedagogical modelling from the beginning teachers' teachers (Calderhead & Robson, 1991). As student teachers' beliefs are based on their individual experiences (Bird, Anderson, Sullivan, & Swidler, 1992), hence, they are robust and they act as a filter through teachers' education programmes are viewed (Wideen et al., 1998). Other researchers have supported this view, indicating that beliefs are deeply rooted and pervasive, and that unless teacher educators engage student teachers to challenge them, attempts to change will be ineffectual (Bryan & Abell, 1999; Tann, 1993).

In a complementary view, Gess-Newsome (1999) established that pre-service teachers construct images of teaching from their previous experiences as elementary and secondary learners, the activity network of the university classroom and their own particular networks according to their culture. Similar findings have been reported in second-career teachers (Tigchelaar, Brouwer, & Vermunt, 2010). According to Davis et al. (2006), teachers' past experiences affect their conceptions of science and their likelihood of specialising in science teaching. The problem mentioned by Smith (2000) is student teachers have typically reported negative experiences such as shame or embarrassment when they were learners, because they were unable to understand science. This can affect their will to teach science.

Nevertheless, other studies have found contradictions between teachers' beliefs and their practices (Ferguson, 2008; Vermunt & Endedijk, 2011). For instance in the research carried out by Duit and Treagust (2003), although some teachers held constructivist views about teaching, their practices indicated their real view of learning seemed to be as a transmission rather than a construction. Simmons et al. (1999), signalled that although teachers professed student-centered beliefs, they behaved in classrooms in teacher-centered ways. Additionally, it has been stated that teachers' conceptions of science do not necessarily influence their practice in classroom (Lederman, 1999; Lederman & Zeidler, 1987). In other words, the relationship between conceptions and teacher behaviour in classroom is not always straightforward (Meirink, Meijer, Verloop, & Bergen, 2009).

Characterising beliefs about teaching, Richardson (1996) found that some pre-service teachers viewed schooling was something in which teachers should hand out knowledge that students memorize, and their teaching was a simple and mechanical transference of information or knowledge. Similar findings were presented by Geddis (1993), in a study with beginning teachers where they evidenced simplistic and poorly articulated images of the teaching role and teacher-pupil interactions.

More specifically, a few studies were found describing or characterising teachers' beliefs about the quality of teaching. Kane and Temple (1997) described in-service teachers' beliefs about a good educator, including elements of character, skills, mastery of subject-matter knowledge, commitment to mission, among others. Murphy et al. (2004) found that pre-service and in-service teachers assigned importance to being caring, patient, not boring and polite. In Weinstein's study (1989), teachers mentioned more frequently items linked with caring, understanding, warmth, the ability to relate to and motivate children and patience. Another study concluded that pre-service teachers thought teacher's personality was more important than teacher's cognitive skills, pedagogical or subject matter, which might reinforce the cultural myth that teachers are born, not made (Surgue, 1996). This valuing of interpersonal aspects of teaching instead of academic goals of schooling could be understood because pre-service teachers sense of good teaching is often defined by their narrow classroom experiences (Woolfolk-Hoy & Murphy, 2001). Nonetheless, a study found that pre-service teachers considered as good quality teaching explaining subjects clearly, enjoying teaching and having knowledge in the subject matter (Strickland, Page, & Page, 1986). This was the only study found where explanations appear to have a role in teachers' thought about quality teaching.

Since the 90's, studies on learning to teach have focused on changing beginning teachers' beliefs to allow them to teach in a different way from how they were taught (Wideen et al., 1998). However, attempts from initial teacher education to change beliefs about quality of teaching are poorly addressed (Murphy et al., 2004). For example, in Tabachnick and Zeicher's study (1999) teaching experience did not impact pre-service teacher's concepts of good teaching , actually it strengthened their prior beliefs only. Another study concluded the same (Murphy et al., 2004). Teachers became more skilful at defending the perspective or beliefs they already possessed (Graber, 1996; Wideen et al., 1998). Indeed, Holt-Reynolds (1992) found that common sense theories prevailed even though pre-service teachers were presented with knowledge-based theories that were contrary to their beliefs.

In another study pre-service teachers referred to their own experiences as learners as their guides for good teaching, implying that they did not shift their beliefs at all during the teacher education programme (Yerrick, Doster, Nugent, Parke, & Crawley, 2003). The same was found by Eder (2005) in Argentinean teachers and Chilean teachers (Latorre, 2003).

This might lead to teachers teaching as they were taught and as they thought they had learned science (Simmons et al., 1999; Trumbull & Kerr, 1993). Thus, it seems very difficult to change beliefs through teacher education or measure a change.

Regarding how to measure the impact of teacher education programmes on pre-service teachers' beliefs, a few ways have been reported, differing in the criteria used to measure it (Wideen et al., 1998). In Garber's study (1996) an initial teacher education programme was described as "having a detectable and substantial influence on the beliefs and actions of program graduates" (p.463). However, the measurements supporting this assertion were not indicated and the paper concludes that it was not possible to identify which elements of the programme had the greatest impact on the mentioned change.

Fang (1996) stated difficulties in measuring teachers' thought processes, due to using techniques such as self-reports which was not enough to catch the complexity of thought processes. Likewise, it has been stated that research in pre-service teachers' beliefs is limited by its dependency on survey methodology (Brookhart & Freeman, 1992). These findings illustrate the difficulties of measuring the influences of teacher education programmes, and also the importance of designing suitable assessment instruments for pre-service teachers' beliefs (Wideen et al., 1998).

A different perspective was given by Calderhead & Robson (1991), who suggested instead of changing beliefs the clue to improve teachers' beliefs about teaching is building on the beliefs that already exist. Similarly, Crawford (1999) asserted it is possible to develop more sophisticated conceptions about science teaching over the course of a year in teachers, and at the same time to align their practice with their conceptions (Bryan & Abell, 1999; Sweeney et al., 2001).

In sum, although many researchers have assumed for long time that prior beliefs of pre-service teachers are very difficult or even impossible to change (Lortie, 1975, 2002; Murphy et al., 2004; Rodriguez, 2001), nowadays this idea has been challenged (Bandura, 1995; Wideen et al., 1998), considering the studies in teacher cognition, teacher thinking and implicit theories which are underpinning beliefs or conceptions.

2.5.2. Teachers' implicit theories

Implicit theories have been described as a system of thoughts that are not clearly articulated or codified by their owners -because of their implicit character- but they are typically inferred and reconstructed by researchers on teacher thinking (Catalán, 2010; Clark, 1988). A similar conceptualization has highlighted their individual character, being named subjective theories. This concept refers to the theories or hypotheses that individuals elaborate (implicitly or explicitly) to make sense of their environment and be able to act in it. According to Catalán (2010), these theories could be also idiosyncratic of a group or community because they are constructed by collective experiences. They have an important function on intergroup relations, mediating the construction of social meaning (A. J. Rodríguez, 1993) and they have a regulatory effect on action (Karmiloff-Smith & Inhelder, 1974; Rodrigo, 1985). The origin of implicit theories was established by cognitive psychology, as the product of implicit or informal learning and the creation of regularities in the world, in order to make it more predictable and controllable (Pozo & Gómez, 1998). For Krause (2005) they are representations that make connections or associations between information units. According to Pozo and Gómez (1998), subjective theories are constructed from a network of beliefs based on implicit suppositions. These assumptions work as a filter of information and help making relations or connections between the information units.

Although the subjective theories could be considered as particular types of beliefs according to Catalán (2010), Pozo and Gómez (1998) remarked they are deeper, more implicit, more stable and more difficult to change than beliefs. This might be because implicit theories tend to be eclectic aggregations of cause-effect propositions from many sources, rules of thumb, and generalizations drawn from personal experience, values, biases and prejudices (Clark, 1988).

Research has looked for teacher theories about their students, the subject matter they are teaching, their roles and responsibilities, how they should act in classrooms, among others (Clark, 1988; Dweck, Chiu, & Hong, 1995; Pozo & Gomez, 2005; Stofflett, 1994; Zanting, Verloop, & Vermunt, 2003). Although the most studied teacher implicit theories are about the students' intelligence (Southerland & Gess-Newsome, 1999), teachers' implicit theories about their work are crucial to be investigated because they play a role in teachers' every day judgements, behaviours and interpretations (Clark, 1988).

The study of teacher implicit theories has used various methods to access teacher thinking, including stimulated recall interviews, linguistic analysis of teacher talk, paragraph completion tests, responses to simulation materials such as vignettes describing hypothetical students or classroom situations, concept generation, group discussions, and mapping exercises to mention the most important ones (Catalán, 2010; Clark, 1988). In a similar route, it has been stated that teachers are not used to articulate their knowledge of practice, and as a consequence, they usually know more than they can say about what they do. This implies there is tacit knowledge which includes reasons for approaching teaching in particular ways, knowledge of teaching procedures and their impact on students' learning (Berry & Loughran, 2010). These authors stated the implicit nature of this kind of teachers' knowledge as the cause of why this is not obvious for them. Then, as they are an economical cognitive way for managing complexity, they are resistant (but not impossible) to change (Pozo & Gómez, 1998).

2.5.3. Modification of implicit theories

To understand how teacher theories can change, it is necessary to understand how they are organized. Catalán (2010) proposed that the content of a theory could be divided into explicit (which was possible to be verbally expressed) and implicit, which is subdivided into implicit content able to be explicit (through inferences from peoples' discourse) and implicit content unable to be explicit.

As is shown by Pozo, Gomez, and Sanz (1999) in the surface level of representational analysis there are the beliefs, conceptions, predictions, judgements and interpretations that people enact to face the situations or task. This level is more accessible and explicit for the person because it is in a more conscious level of representation. Usually conceptions are activated in a specific situation and most of time they are constructed responding to the contextual requirements. They are not necessarily held permanently in the cognitive system because of their contextually specific and unstable character.

The changing theories process requires a deep restructuring in the implicit suppositions, conducting a conceptual change to overcome the restrictions imposed by the person's cognitive system. This change should operate in the deepest conceptual structures, restructuring them in order to construct new knowledge (Pozo & Gómez, 1998). According to Karmiloff-Smith (1992), specific level of representation should be re-described in new and more complex categories in a sequence of progressive complexity in order to integrate or re-interpret previous ideas in others.

Taking into account the findings about distributed memory, it is possible to assume that different implicit theories can be activated at the same time according to the specific context (Pozo et al., 1999). In the case of teachers' change of theories, as they constitute a very cost -effective way of reasoning, the theories need to be addressed and confronted with practice (Pozo & Gómez, 1998). Besides, beliefs or conceptions need this process to be modified: restructuring, making them explicit and integrating (Pozo et al., 1999):

(a) The restructuring process implies to look for different suppositions, attributions, generalizations or abstractions in people's thinking.

(b) The progressive making explicit process adds conscious about differences between the prior knowledge and possible new theories through studying concrete situations. In this process people go deeper in their own representations to make them explicit in order to modify their theories' foundation.

(c) The hierarchical integration process implies that the elemental theories are integrated or re-described in more complex ones. It needs a metacognitive effort in order to construct theories with a more complex structure and with a higher explicative power.

Thus, the imperative for science teacher educators would be to make the implicit knowledge more explicit (Southerland & Gess-Newsome, 1999), it means making theories progressively fit into a position where they can be affected. Recent studies have proved that implicit theories can change through this restructuring process (Pozo & Gomez, 2005), which was previously seen as an articulation of conceptions, implicit theories and principles of practice in a reflexive process (Clark, 1988). Concept maps, metaphors and flow charts, are techniques to aid pre-service teachers in the elucidation of their situated position. Also, using the same input twice, once at the outset of the course and again at the end of it, offers to pre-service teachers the possibility to look for transformation and change (Ferguson, 2008).

It is interesting to note that in programmes oriented to work with students, specifically trying to elicit their thoughts in a group in which positive results were reported, group size appeared to be a factor; small numbers of participants often were involved, and they often worked in groups in a close relationship with the instructors. It has been suggested that innovative programmes provide a form of shelter for students within which they can examine and assess their beliefs among peers (Wideen et al., 1998).

In sum, a goal of teacher education should be to help pre-service teachers challenge and refine their ideas about teaching and learn how to learn from their own teaching experience. By understanding their conceptions, practice and the relation between both it is possible to create supportive environments for pre-service teachers' development (Bryan & Abell, 1999). This might lead to design effective teacher education programmes which recognise the development of teachers' knowledge as neccesary (De Jong et al., 2005). In this current research, implicit theories about explanations in science were investigated. In order to understand the important role of explanations in science teaching, details are provided in the next section.

2.6. Explanations in science teaching

School science teaching needs to address several and different goals. Some authors like Arzi and White (2008) put the emphasis in the subject matter knowledge, and others in the skills pupils need to develop to create science (Akerson & Hanuscin, 2007; Eick & Reed, 2002; Windschitl, Thompson, & Braaten, 2008). However much of the real work in science classroom appears as describing, labelling or defining concepts (Ogborn et al., 1996), which might support why explanations are considered to be at the centre of science education (Geelan, 2009). The approaches and strategies in science teaching are described below, to situate explanations in their broader context.

2.6.1. Approaches in school science teaching

Different approaches of school science teaching have been described highlighting teachers' role:

(a) Traditional: It is focused on verbal knowledge transmission. The teacher provides the knowledge and regulates students' learning (Vermunt & Endedijk, 2011).

(b) Expositive: The teacher defines the knowledge to the students in a verbal exposition which develops students' ideas of the scientific concepts which are seen as the core of science curricula (Arzi & White, 2008; Pozo & Gómez, 1998).

(c) Finding out: Teachers need to guide students to research and reconstruct scientific findings (Windschitl et al., 2008).

(d) Cognitive conflict: Teachers need to confront students' conceptions in conflict situations in order to get a conceptual change and substitution of old theories for more scientifically oriented models (Druit, 1999; Duit & Treagust, 2003).

(e) Inquiry: As mentioned in section 2.2.1, it has been widely recommended in science education reforms. Teachers need to guide students' construction of theories and models about science and transform students' observations into meaningful knowledge (Davis et al., 2006). Teachers must change their participation in the activities according to students' needs to scaffold the students' learning process (Pozo & Gómez, 1998).

2.6.2. Science teaching strategies used in classrooms

There are strategies described mainly for primary science teaching. For example Eshach (2006) suggested familiarization of new scientific terms by announcing part of the term, referring intentionally to a wrong possibility (to introduce concept application in an incorrect or absurd situation), and using similar concepts present in real life. These are useful to clear scientific ideas in students' minds through comparison, differentiation, etc. Also, this researcher stated explaining new terms' verbal meaning to primary learners, moving around the terms to mediate real understanding as important for science teaching. Other plausible strategies to carry out is making the pupils develop science research projects and share them with classmates (Ward, Roden, Hewlett, & Foreman, 2008; Wellington & Ireson, 2008) or underlining the difference between what pupils already know and what they are going to know. When the tension is established, there is a gap to be bridged with understanding. In this gap, it is possible to construct explanations in science teaching as stories (Ward et al., 2008) invoking the "protagonists" or "entities" that are the scientific concepts (Ogborn et al., 1996). Finally, other strategies are role playing, activities outside the classrooms, encouraging pupils' conversations about scientific topics or receiving visits from people who work in science professions (Braund, 2008; Braund & Reiss, 2004). For secondary science teaching there are recommended strategies to advance pupils' scientific reasoning, allowing them to observe, hypothesize, use appropriate apparatus, measure, interpret data and draw generalizations (Eshach, 2006).

Strategies suggested independently on the student age are: explaining (Geelan, 2012; Ogborn et al., 1996), using analogies and/or metaphors (B. González & Moreno, 1998; James & Scharmann, 2007), demonstrations or experimentations (Ogborn & Martins, 1996), introducing new topics stressing its usefulness or writing it down, using a gesture, a diagram or chart, changing in the pace of speech or voice intonation, the repetition of an idea or making seem strange the familiar or comfortable things (Ogborn et al., 1996). Gestures in science teaching can express new levels of understanding and new concepts (Roth & Tobin, 2001; Roth & Welzel, 2001).

Osborne & Freyberg (1985) indicated several years ago that using combined strategies (for example demonstration and analogies) improve science teaching intelligibility and plausibility. Nonetheless, at the moment there is a limited range of these strategies that are really used in classrooms (Osborn & Dillon, 2008). In Chile for instance, the most frequent strategy that teachers use in science classrooms is the conceptual explanation (Preiss et al., 2012). Explanations in general and particularly in science teaching are described in the next pages.

2.6.3. Context and definition of explanations

An explanation is basically seen as an act intended to make a phenomenon clear, understandable or intelligible for others (Brewer, Chinn, & Samarapungavan, 2000; Danto, 1985; J. Kim, 1995; Norris, Guilbert, Smith, Hakimelahi, & Phillips, 2005). The classic view of explanation defined it as an answer to a why-question (Norris et al., 2005), with the purpose of sharing knowledge and meaning (Treagust & Harrison, 1999). However, this definition could imply a similarity between explanation and description, which may be confusing. Descriptions are “pure information, isolated and without a network of relatedness” and explanations contain “information with connections, a relationship built on a system of causality” (Edgington, 1997, p. 41).

Another similar definition asserted explanations are systematic arguments that address the issues of how and why a phenomenon happens, usually including cause and effect statements, while descriptions are statements focused on superficial details only (Treagust & Harrison, 1999). Even so, teachers and students have used description and explanation of an event or process as equivalent, interchangeably (Horwood, 1988).

At the philosophical level there is debate as to whether an explanation is a process -the act of explaining- or just the syntactic product. If considered as a product, the explanation would be isolated from the explainer, the audience and the context (Treagust & Harrison, 1999). In this research the explanation is considered as a product and process, because it is likely that both the product and the process of explaining are equally relevant for helping students to understand scientific phenomena. Besides, taking into account the contextual variables involved in the teaching process might orientate towards improvements. ‘Explaining scientific phenomena to school students involves both process and product, because an explanations’ viability is determined by its context’ (Treagust & Harrison, 1999, p. 32).

2.6.4. Scientific explanation vs. science teacher explanation

Explanations in science have received attention not only from the education field, but also from the scientific field and from philosophy (Edgington, 1997; Sevian & Gonsalves, 2008), especially regarding what defines a scientific explanation (Geelan, 2012). Scientific explanations and science teacher explanations are described as follows.

Scientific explanations are formulated as deductive arguments whose conclusion is the explanandum sentence, thus, they are the answers to the question why does a phenomenon occur? (Edgington, 1997). The essence of scientific explanation is communicating understanding about a phenomenon under investigation to the scientific community, they are evidence-driven and use correct scientific terminology (Treagust & Harrison, 1999). Norris et al. (2005) differentiated characteristics of scientific explanations according to their function. These are summarised in Table 1. Sevian and Gonsalves (2008) described that the common framework used by scientists to explain phenomena are causal, functional and intentional.

Table 1: Scientific explanations' characteristics by function

(Adapted from Norris et al. (2005, p. 550))

Explanatory function or type	Characteristics
Interpretive explanation	It clarifies meaning, defines terms, propositions, treatises and assigns, develops or expands meaning.
Justificatory explanation	It justifies why something was done, provides reasons for acting, appeals to norms, standards or values and it may appeals to causes as reasons for acting.
Descriptive explanation	It describes a process or structure.
Causal explanation	It cites a cause for events or laws.
Deductive-nomological explanation	It explains particular facts by deriving them from general laws and other facts. It includes at least one universal law. Its basic structure is a deductive argument.
Statistical explanation	It explains showing facts to be highly probable. Its basic structure is an inductive argument. It includes at least one statistical law. Causation typically is not implied.
Functional explanation	It explains facts by indicating their function.
Explanatory unification	It explains phenomena by fitting them into a general world view, aims to derive facts from smallest number of assumptions. It views ideal explanations as deductive.
Pragmatic explanation	It explains by answering why questions. Questions are asked and answers are given in a context which enables determination of appropriate contrast classes and relations.
Narrative explanation	It narrates events leading up to its occurrence. It cites unique events as explanatory of other unique events. It seeks unification. It rarely supports predictions, but relies upon retrodiction to indicate the present is consequence of past.

The idea of explanation as an answer to a why-question has been enriched by the construct of an explanatory framework. This is the way in which teachers use analogy, metaphor, examples, axioms and concepts linking them together into a coherent whole for the classroom (Geelan, 2003). Explanations of natural phenomena are used every day for delivery and assessment of instructions in science, in lecture demonstrations, in science classrooms, in educational and curricular materials, and tasks used for assessment of students' understanding (Edgington, 1997). Eder (2005) described classroom explanations as a didactical strategy that constitutes the heart of every teaching episode. A teaching strategy was defined by Canal de Leon (2000) as an intentional and conscious guide that gives a general regulation to the teaching activity and provides sense and coordination to everything done to achieve a goal. Having this definition, it is possible to state that teacher explanations are not necessarily antithetical to inquiry learning or other types of constructivist understanding of teaching or learning, and also they are not restricted to lecturing or expositive teaching only (Geelan, 2012). Indeed, teacher explanations express implicit messages about the nature of science and they might promote curriculum goals (Edgington, 1997), and usually, explanations are collaboratively examined and generated in the classrooms, constructed from fragments of students' and teachers' ideas (Dawes, 2004).

Besides, science teacher explanations are explanations for someone to learn (Carr et al., 1994; Horwood, 1988). This difference was remarked by Treagust and Harrison (1999), pointing out that science explanations are driven by experts and delivered to expert audiences, while science teacher explanations are given to novices, usually primary or secondary students that do not know the meaning of scientific terms. Science teaching explanations differ from scientific explanations in rigour, length and detail (Treagust & Harrison, 1999), but not in importance (Geelan, 2012). In fact, for the students a good explanation makes a difference to what counts as a phenomenon, because it tells them what is relevant to look for. Further, transformation of scientific knowledge into a carefully versioned form is the most important aspect that defines a teacher explanation (Ogborn et al., 1996).

An interesting study conducted in Argentina by Eder (2005) asked pre-service teachers about the difference between scientific explanations and science teacher explanations. She found that for pre-service teachers a scientific explanation was circulated among peers who could discuss it, prove it or accept it, while science teacher explanations were driven in a hierarchical context because the teacher knew what was correct and the students did not. Then, teacher should move students to the correct scientific meaning through the explanation. Furthermore, pre-service teachers assumed that as scientific explanation operates with abstract entities, it requires an audience with logical thinking, while teacher explanation needs concrete examples and a shared language that allow communication and understanding between teachers and pupils.

2.6.5. Styles of teacher explanations

Even though teacher explanations have been seen as a device to share knowledge and meaning to those who do not have prior or sophisticated knowledge and understanding of the phenomenon or concept in study (Treagust & Harrison, 1999), their purpose is not only sharing knowledge. In this research is understood an explanation transforms scientific knowledge continually, in such a way to be memorable, intelligible and able to be put into use in the school, making it accessible to pupils (Ogborn et al., 1996). There are different explanation's styles considering the interaction between students and teacher, as presented in Table 2.

Table 2: Teacher explanation style

(Adapted from Ogborn et al. (1996)).

Style proposed	Description of the style
Thinking together	The teacher describes the ideas emergent from the students, connecting their ideas with scientific models after.
Telling a story	The teacher turns the explanation into a narrative, a type of tale that integrates the different points of view, visions and concepts, opening up opportunities for contributions from the pupils, and reworking the ideas that have been obtained.
Saying it or seeing it in teachers' way	The teacher asks to pupils to re-describe their ideas, to reinterpret these ideas in other scientific model terms given by the teacher, encouraging pupils to use more precise language.

In the current research teacher explanation is not seen as a monologue from the teacher but as a dialogue, where the teacher creates diverse explicative settings contrasting them with the students, and where students can also explain scientific ideas to other students (Dawes, 2004). In fact, explaining science to the students with the students is one of the most challenging tasks for teachers (Ogborn et al., 1996), although other types of explanations have been described. These are presented in the next pages.

2.6.6. Types of teacher explanations

Dagher and Cossman (1992) presented a classification in types of verbal explanations in a study with twenty in-service teachers. Teachers were working in middle public schools and their classes were observed, recorded and transcribed verbatim. The classification was done using the constant comparative method (Strauss, 1987), and ten types were described and exemplified, as shown in Table 3 (Adapted from Dagher and Cossman (1992, pp. 364-366)).

Table 3: Types of teacher explanations

Type of explanation	Description of the type
Analogical	A familiar situation, similar to the unfamiliar phenomenon to be explained is used to provide the explanation.
Anthropomorphic	A phenomenon is rendered more familiar by attributing human characteristics
Functional	The phenomenon is explained in terms of its immediate consequence or function.
Genetic	The explanation is provided by relating an antecedent sequence of events, and the focus is about what happens, not why it happens.
Mechanical	Causal relationships are given to explain the phenomenon.
Metaphysical	A supernatural agent is identified as the cause of the phenomenon explained.
Practical (How to):	It involves instructions as to how to perform physical or mental operations or procedures.
Rational	During the explanation, evidence or warrant is provided for a given claim in an effort (implicit-explicit) to compel belief. These explanations involve giving support to a claim that has been made in order to compel others to accept that claim.
Tautological	The how or why question or statement is reformulated in the explanation without adding any new information to its content.
Teleological	A phenomenon is explained in terms of how its immediate consequence (function) contributes, through concerted action with other phenomena to the probable attainment of an ultimate consequence (goal).

Even though these types are ideals, Dagher and Cossman's study (1992) found that most of the explanations given by the science teachers in middle-schools were genetic, mechanical, practical and analogical, in this order of frequency. The teachers who provided more explanations did not necessarily use a greater variety of types.

When teachers decide to explain a scientific concept in a particular way, science content plays an important role in framing teachers' explanation: firstly depending on the background knowledge the teachers had (Faye, 2009) and secondly it depends on certain characteristics of the content. Actually the genesis of the nature of science teacher explanations may be explained by the range of properties of the content present in the explanation (Treagust & Harrison, 1999). Regarding the last point, White (1994) described ten properties of science content that might influence the type of teachers explanations which are exemplified in the Table 4 (Adapted from White (1994)). For example, as atoms and magnetic fields are non-observable and abstract concepts, they are often presented in models and images.

Table 4: Properties of contents that influence teaching

Property	Description	Examples
1. Openness to common experience	The extent to which the content is present in common experiences and thus, is posing prior beliefs in people about it.	Force and light (common-has many experience conceptions) compared to atoms (uncommon-has fewer alternative conceptions).
2. Abstraction	How tangible or abstract is the content in a continuum.	Speed is common and tangible, while acceleration is abstract.
3. Complexity	It is about the number of elements that it comprises and the coherence between them.	'Density' involves only mass and volume, and the effects of temperature, but 'sound' includes many contributory concepts.
4. Presence of alternative models of good explanatory power	Whether there are alternative models of the concept or not and how they can explain other related concepts.	The model of 'heat' as associated with the kinetic energy of models displaced a model of it as a fluid, caloric. Then, caloric view of heat has a good explanatory power.
5. Presence of common words	It refers to availability of the words that are part of a concept in non-scientific language.	Words like animal, flower force and work have different meanings in everyday and scientific contexts.
6. Mix of types of knowledge	Knowledge can be iconic propositional, procedural, etc. A concept might be compound of several or just one type.	Propositions, images, analogies, episodes, and procedures. When teacher should use each of them to be sure students process each type of knowledge?
7. Demonstrable vs. arbitrary	It refers to how demonstrable is the concept during teaching.	Differences between flowering and non-flowering plants are demonstrable, but the statement that current flows from the positive terminal of a cell is arbitrary.
8. Social acceptance	How accepted is the concept in society? It implies how easily debatable the concepts are.	Creation and evolution, population control, power generation are contentions.
9. Extent of links	It refers to the extent to which a topic can be related with other content, or how self-contained they are.	Energy is a key topic in physics, biology, etc., as, to a lesser extent, is electricity. Pervasive topics might be taught differently from restricted ones.
10. Emotive power (interest)	Different topics are likely to arouse different types and intensity of emotions in people.	There might be more potential wonder in sinking and floating than in rusting.

Although the styles of explanations may have different emphasis, they all are mediated by teacher's speech. This is the reason why this research focused in verbal explanations of scientific concepts as a common point. Explanations are compound by different elements. Some of them were considered in the research rubric and are described in the next pages.

2.6.7. Explanation’s elements (rationale of research rubric)

For the explanations criteria definition, elements from the literature review were selected and organized. Some of these were specific for science education and others could apply to any teacher explanation of concepts. Other elements were not suitable to be incorporated in the rubric; elements that were beyond the explanations and useful in the whole lesson were too wide to be part of a specific instrument to assess explanations, such as questions posed by the teacher or answers to pupils’ questions (Carlsen, 1993), gaining students’ attention or motivation (Mohan, 2007). The criteria and elements described as follows constituted the rationale of the research rubric used to assess the quality of the student teachers’ explanations. The first six (A) referred to the structure of the explanation, and the last four (B) characterized the complementary ways which the teacher could deliver it.

A. Structure describes the logic and flow of an explanation (Sevian & Gonsalves, 2008).

Criterion 1: Clarity

An explanation requires pointing out features, patterns and structuring the content in a clear and focused manner (Sevian & Gonsalves, 2008). It means the language the teacher uses is adequate and understandable for the pupils and the concepts are presented in simple words (Danielson, 2011; Geelan, 2009; Gobierno de Chile, 2003; Sevian & Gonsalves, 2008). Also, the teacher slows the word flow down when ideas are difficult or complex to understand (Mohan, 2007). This element was part of the rubric.

Another aspect of clarity is reflected in the teachers’ specialized language usage in school science. It becomes acute when the focus is on mere use of technical terms. To avoid this problem, teachers should make a separation between a description of the phenomena and the explanation of why it happened that usually involves a scientific concept (Wenham, 2005). Then, it is suggested teachers should first describe and then explain, because unlike description, scientific explanation does require the use of special concepts and scientific language. Actually Wenham (2005) indicated that technical terms should be used only after the explanation has been developed to communicate what has been found out. This shows clearly the role of scientific concepts in the explanation and was considered as the second element in the clarity criterion. Finally, an important element related to clarity was signalled by Geelan (2012) who stressed the importance of avoiding tautology in the explanation.

Criterion 2: Coherence and cohesion

Windschitl et al. (2008) underlined that explanations are not only about patterns in observable relationships but about how these relationships act as evidence for why a phenomenon happens in a particular way. Coherence is the criterion about how the text knits together, making meaning (Halliday & Hasan, 1976). Rodrigues (2010) asserted it could be achieved through establishing relations of cause-consequence between the ideas in the teachers' talk, inclusion or exclusion, differentiation or similarity, among others. This was the operationalized indicator in the rubric.

Likewise, cohesion could be determined by considering the ties or links in talk that internally relate clauses, parts and sentences cohesively. They are important in teachers' talk because if the links are weak, then the listener may experience difficulties in discovering the speaker's intention (Rodrigues, 2010). The conditionals could define the logical steps in an explanation (Sevian & Gonsalves, 2008). Then, including strong ties in the explanation should be taken as an important element in the rubric.

Criterion 3: Sequence

An essential aspect in effective science teaching is the organization of the elements of an explanation (Mayer & Jackson, 2005). Experts structure knowledge according to organizing principles of the discipline, and an effective explanation requires articulating those organizer principles and conditions associated with the concept (Sevian & Gonsalves, 2008). This is usually called the sequence.

In terms of structure van Peer and Chatman (2001) remarked that every explanation should have a logical structure of a formal argument, and it could be seen as a recounting of events structured in time. What is being explained should be logically deducible from the antecedent conditions and general laws (Norris et al., 2005). Also in the Chilean context it has been mentioned that a good explanation is structured in a logical and consistent sequence (Gobierno de Chile, 2003). Hence, teachers should dissect the concept into understandable parts, joining them progressively and using each time more abstract language to facilitate the students' identification of the concepts' characteristics (Sanmartí, 2000). The progressive character of the sequence was also indicated by the Chilean national framework for good teaching, which has remarked that the sequence should progress in a coherent manner and facilitate the students' understanding. As a consequence, the explanation would be a unified globalism (Gobierno de Chile, 2003), going for example from the simpler to the more complex aspects of the scientific concept being taught (Sevian & Gonsalves, 2008), or triggering the related material contiguously in space and time (Wu & Shah, 2004) to facilitate the connections between the elements (Cook, 2006). Thus, progression was an element incorporated in the rubric.

Also, the teachers should scaffold their instruction in order to enhance the students' cognitive development (Appleton, 2007) and reduce the cognitive load (Cook, 2006). In the scaffolding process each part of the explanation should play an important role in the totality of the explanation to make it coherent (Gobierno de Chile, 2003). Otherwise, if the facts are disjointed and not held together by articulated guiding principles, the explanation could become a string of unconnected facts (Sevian & Gonsalves, 2008). This was the second element of this criterion.

Criterion 4: Accuracy

Accuracy is the extent to which the explanation is scientifically correct in terms of factual knowledge (Sevian & Gonsalves, 2008). Teachers should express in their explanations a deep understanding about the specific body of content knowledge (Gobierno de Chile, 2003). Gess-Newsome (1999) also stated the importance of content knowledge for teaching science. Likewise, according to Danielson (2011) teachers must understand the central concepts, tools of inquiry and structures of the discipline they are going to teach, to be able to create learning experiences that make the discipline accessible and meaningful for learners to assure mastery of the content. This could be observed when scientific terminology was correctly used by teachers (Sevian & Gonsalves, 2008), and the variations of this -such as inaccuracies or content errors (Danielson, 2011) -were the elements incorporated in the rubric.

Criterion 5: Sufficiency

Another interesting aspect to evaluate in science teachers' explanations is whether the globalism that the teacher develops in the classroom contains or not the main aspects that contribute to the concept's construction. Roth & Welzel (2001) mentioned as an analysis category if teachers' talk in and of itself is sufficient or insufficient to understand just what the explainer wanted to explain. The conceptual explanations have a teaching objective, according to how deep the teacher wants to go. In this sense, it is valuable that the teacher covers thoroughly the concepts they propose to teach (Danielson, 2011), respecting the students who need going slowly in their learning or having successive views to understand (Cook, 2006).

Criterion 6: Connection with students' experience

Building up the explanations on students' prior understandings, ideas, and knowledge has been emphasized as necessary in order to connect the explanation with a familiar context for pupils (Sevian & Gonsalves, 2008). Teachers should explain the meaning of scientific concepts trying to connect it with previous students' ideas or concepts that they bring to the lesson (Eshach, 2006; Faye, 2009; Ogborn et al., 1996; Treagust & Harrison, 1999). Consequently, Marzano et al. (2001) suggested that asking questions to elicit prior understanding before presenting new content is crucial. It enhances information retention and facilitates the integration of new knowledge with prior knowledge (Cook, 2006). This was an element considered in the rubric. Furthermore, what teachers do with the ideas or everyday knowledge elicited is also important. According to Smith (2000) pre-service teachers need to respond to children's naïve ideas, and Limon & Carretero (1997) mentioned that introducing cognitive conflict in the students' prior ideas and scientific conceptual ideas is a clear way to initiate change in students' conceptions. Also, Pozo and Gómez (1998) signalled the importance of the contrast between the scientific arguments derived from the teacher and the pupils' arguments. Gil (1994) remarked on stating explicit relations between the students' prior ideas and the scientific ideas, and Sanmartí (2000) stated that establishing good connectors between everyday life and scientific knowledge is important for their differentiation. Sanmartí (2000) also indicated that the connection with students' prior learning is a way of making the concepts meaningful for them. The connection of scientific concepts with the students' everyday life allows them to understand more abstract ideas in a known context and to concretize other ideas that are difficult to imagine. Hence, this is another element included in the rubric.

B. Complements are strategies to support the explanation's presentation (Mohan, 2007).

Criterion 7: Metaphor, analogy, simulation or model usage

Significant research attention has been paid to the use of analogies in teaching science. This work forms the largest single body of literature in relation to explanation in science education (Geelan, 2012). Mental imagery is a powerful learning tool (Bellezza, 1996; James & Scharmann, 2007) and using models and analogies could promote conceptual change in pupils' prior ideas (Limon & Carretero, 1997). Analogies can stimulate new inferences and insights, and advance the conceptual understanding of scientific phenomena (Glynn, Taasoobshirazi, & Fowler, 2007; B. González & Moreno, 1998; Wong, 1993). They are especially useful to teach difficult or abstract concepts (Podolefsky, 2007) because they bring content to life (Danielson, 2011). Actually it has been proved that students who were taught with analogies got better results in science learning tests (Dupin & Joshua, 1989).

Metaphors, models and simulations could encourage the creation of mental images that scaffold the explanation (Geelan, 2009; Sevian & Gonsalves, 2008). This point is also highlighted by Sanmartí (2000) and Dagher (1992), who indicated that analogies and metaphors are central in the evolution of scientific learning. Likewise, Zacharia (2005) signalled that when teachers interacted with simulations or models of a scientific phenomenon, the explanations that they constructed were richer, more detailed, scientifically more accurate and involved more formal reasoning. Hence, using correctly a metaphor, analogy, model or simulation was considered as a rubric element. The utility of analogies has been seen in reinforcement of what students already know, progressive transformation of their ideas, encouraging pupils' imaginative potential and conceptual flexibility, and refinement of teachers' explanations (Butefish, 1990).

Nevertheless, for success in using these devices it is crucial to develop explicitly the connection between the analogy and the scientific target situation and to dialogue about the analogy with the students, rather than simply presenting it (Brown & Clement, 1989 as cited in Geelan, 2012). The pupils need to identify the concepts in teacher-constructed analogies (James & Scharmann, 2007). Also, it is important to connect the metaphors, analogies, simulations or models characteristics with the definitions and the science content (Sevian & Gonsalves, 2008). This important aspect was incorporated in the rubric.

Criterion 8: Example, demonstration, experiment, graph or image usage

Verbal explanations in science need to be complemented by visual and tactile representations (Cook, 2006; Geelan, 2009, 2012), indeed, generating and using evidence as support for the explanation is central (Windschitl et al., 2008). Demonstration helps students to create representations about science (Sevian & Gonsalves, 2008) and using examples, images or graphs students could create mental images. Also, demonstrations and experiments put science meaning into matter. Scientific theories talk about a world behind appearances, and demonstrations try to bring that underlying world to the surface. Their objective is to persuade the student that the things are as they are shown by the teacher (Ogborn et al., 1996). Specifically in science teaching, visual representations provide a means for making visible phenomena that are too small, large, fast, abstract, invisible or slow to see with the unaided eye or recall from direct experience (Buckley, 2000). Graphics are used in science to display and organize information, and promote a shared understanding of scientific phenomena (Kozma, 2003), presenting relationships and processes difficult to describe orally (Cook, 2006).

These representations benefit the pupils' understanding because they provide resources for how to think about the scientific phenomena (Ogborn et al., 1996). Likewise, they can be used by the teachers to illustrate some aspects of the concepts and also shared with other students and critiqued. Hence, they can help learners to recognize gaps or insights to be addressed before moving forward in scientific understanding (Windschitl et al., 2008). However, the relevance of showing how the features of the representations interact and interrelate is also relevant (Mayer & Gallini, 1990). It is a problem if the teacher paints the image effectively but does not link it to the science. The representation would not scaffold the explanation as the teacher would not weave together the image as an example to build the understanding of the scientific concept (Sevian & Gonsalves, 2008). Likewise, in experiments and demonstrations, special attention should be paid to the reasons that explain why the demonstrated or experimented phenomenon worked in determined way (Campanario & Moya, 1999). Experts link representations and verbal explanations with the underlying principles of the content to develop a more comprehensive mental model (Snyder, 2000). Hence, the teacher should establish the link. Both elements -supporting the explanation with a representation and linking it with the scientific concept- are in the rubric.

Criterion 9: Gesture and voice usage

The use of voice inflections and body language are considered as part of effective communication methods. In explanations they serve as emphasis variation (Mohan, 2007). Teachers' hand gestures can assist the pupils' understanding if they are connected with teachers' talk and with the content (Geelan, 2009). However, they could also be unhelpful if they are not used carefully or distractive when they are not coherent with what is been said (Sevian & Gonsalves, 2008). A hand or body gesture can represent concept aspects that are more abstract or difficult for the students to picture, and they also can highlight some properties of the concepts (Roth, in print; Roth & Tobin, 2001; Roth & Welzel, 2001). Voice inflections or changes in the pace of the speech can have the same function, emphasizing some points of the explanation and or helping pupils to differentiate the important from the less important parts (Sevian & Gonsalves, 2008). Both elements were considered, taking into account that non-verbal interaction is key to effective instruction (Marzano et al., 2001).

Criterion 10: Misconception illustration

Misconceptions could involve the misunderstanding of factual information, coming from parents and teachers or be constructed based on students' experience (Martin, Sexton, & Gerlovich, 2002). Then, teachers should point out possible areas for misunderstanding and be alert to students' revealed misconceptions (Danielson, 2011). Teachers need to be careful to introduce new topics in such a way as to prevent pupils from developing misconceptions (Thompson & Logue, 2006), and correct them when they appear in pupils' activities (Danielson, 2011). These two elements were the last ones included in the rubric.

2.6.8. Effective science teacher explanations

There are a few theoretical or empirical works regarding science teacher explanations that have mentioned characteristics implying a more effective (or better) explanation. Authors like Mohan (2007) have indicated that gestures, stressing points or linking words could made an explanation more effective . From Feynman's work (1994) it was possible to interpret that expert explainers use their imagination to create devices to make sense of abstract, difficult or non-observable science concepts.

One of the only characteristic agreed as part of effective teacher explanations is their **accommodation** to the explainer, the audience (Leite, Mendoza, & Borsese, 2007), the content and the context where they are created (Carr et al., 1994; Treagust & Harrison, 1999). Unless the explanation is understandable for students, it will not be a useful explanation (Wragg & Brown, 2001). Then, explanations should sensitively accomodate students' features and needs (Treagust & Harrison, 1999). Good explanations' features according to Eder (2005) are their clarity, simplicity, concreteness and containing adequate examples. For instance, an explanation for secondary science should be process-oriented and less dominated by facts than one for primary school (Treagust & Harrison, 1999).

In the conceptualization of Faye (2009) an explanation usually contains a causal element and it is not a circular explanation -where the information explained is a fact explained by the same fact- . Geelan (2012) also stressed the importance of avoiding tautology in the explanation. Otherwise, Ogborn et al. (1996) suggested any explanation should include four components; establishing the difference between what the students know and what they are going to know, constructing entities of the explanation, transforming the student's knowledge and demonstrating the phenomena. Nonetheless, no one of the authors reviewed here mentioned how the explanations were evaluated or how the effectiveness was established.

2.6.9. How to assess the quality of explanations?

This specific question has not been directly addressed in science education. However, from some works it is plausible to interpret or deduce criteria to orient a possible answer.

On one hand, Ogborn et al. (1996) mentioned it is not the goal of explanation forcing students to go to the correct model that is previously established, but promoting pupils' reflection, metacognition and model contrasting. In a similar line, Carr et al. (1994) mentioned there is not a single explanation for phenomena definition, and this should be addressed in science classrooms interactions. On the other hand, there are authors like Faye (2009) who emphasised that the information explanation gives must be correct, because it provides understanding and it gives the psychological feeling of knowing to the students. This view tries to conduct the students' ideas to the scientific model proposed to make sense, more than construct the scientific knowledge with the students to build scientific understanding. Then, it is possible to address here the intention of explanations as an important issue.

In the case of beginning science teachers, a study conducted during the nineties found that their explanations included logical flaws, as well as errors of scientific fact (Goodwin, 1995). This study was giving a prompt about an aspect that could imply a quality characteristic: the accuracy of the facts contained in the explanation, also underlined by Treagust and Harrison (1999). Conversely, another study conducted with pre-service teachers in Portugal, Spain and Italy suggested that the misconceptions found in teachers' explanations were more related to the teachers' lack of content knowledge in the field than with their skills in explaining concepts (Leite et al., 2007). Here, the notion of a set of skills to explain appeared as possible to be assessed or at least separated from the teachers' content knowledge. In a similar line, Mohan (2007) described "explaining" as a skill compound by different sub skills divided into structural aspects (including introduction, key concepts, examples and summary) and presentation aspects (comprising style, clarity, pointers and linkers). In a study about the learnability of explaining, different subject university lecturers expressed through a questionnaire that most of the elements to explain could be learnt, to varying degrees (Brown & Daines, 1981). As it was based on perceptions exploration, it should be taken descriptively only showing the modifiable character attributed to explanations by this group of teachers.

Nevertheless, none of the previous researchers proposed a way to assess whether the explanation is being conducted in one or the other way, or how to measure the different elements of the explanations to assure they behave together as a skill.

The only work found in this literature review oriented to judge the quality of explanations was proposed by Sevian and Gonsalves (2008), but it was not developed in the science teaching or teacher education area. They created a rubric to analyse scientific explanations about their research given by science graduate students. In the instrument construction process, videos and transcripts of the explanations were coded, categorized and evaluated by an expert to identify patterns in the ways in which the graduate students communicated their research to different audiences. Sevian and Gonsalves's rubric contained three categories; PK, CK and PCK understood as the integration of the two first. The researchers described four performance levels for each element; "not apparent", "emerges but inadequately", "present but inconsistent" and "consistent". Within PK, they found four criteria: structure and balance, response to the audience, choice of language and technical skills in using media. Grouped as CK they considered how the knowledge was organised and the students' ability to transfer knowledge to broader contexts. In PCK they measured the usage of mental images that supported the explanations, tactical usage of media and scaffolding or waving examples together.

After applying the rubric, they concluded assessing students' explanations was possible. Giving the students the opportunity to analyse their explanations enabled them to identify strengths and weaknesses in their presentation skills, understanding the relationship between science content and the structure of the explanation, which helped their development of PCK. The researchers recommended recording and transcribing the explanations to assess their structure, analysing the videos holistically first, and then more deeply. Finally, they argued that having a rubric "is not only good for evaluating the effectiveness of a given science explanation, but also for preparing both scientist and science teachers to explain science more effectively" (Sevian & Gonsalves, 2008, pp. 1463-1464). This instrument and their suggestions were an important input for the current research, where a different instrument was constructed but following the same rubric format. This and all others research methodological aspects are described in the next chapter.

3. Methodology

This chapter starts stating the research objectives and the approach this research had. Then, it is moved into the research design, presenting in detail the peer assessment (PA) intervention design, the questions that guided this research, the sampling and participant characteristics. After this, the measurements, instruments and data collection techniques are described, to continue with the procedure that this research followed. Finally, this chapter ends with a review of the data analysis processes used in this research.

3.1. Research aims

The general aim was to explore to what extent peer assessment (PA) could facilitate change in Chilean pre-service science teachers' conceptions and practices to explain scientific concepts during initial teacher education (ITE).

The subsidiary aims were:

(1) To explore the implicit theories about the quality of conceptual explanations in pre-service science teachers.

(2) To determine whether differences existed or not in implicit theories about conceptual explanations according to pre-service teachers' science knowledge.

(3) To analyse and compare the conceptions about the quality of teacher explanations of pre-service science teachers who had been exposed to PA and who had not been exposed to PA.

(4) To determine the quality of pre-service science teachers' conceptual explanations.

(5) To compare the quality of pre-service science teachers' explanations before and after PA.

(6) To identify the main elements associated with the pre-service teachers' process of change from the teachers and researcher's perspective.

(7) To determine if good practices to explain scientific concepts were transferable into real teaching practice.

(8) To identify facilitators and obstacles for teachers' skill transference from ITE to real teaching practice.

3.2. Approach

The model in this research entailed pre-service teachers engaging in peer videoing of simulated class teaching and the subsequent analysis of their practice in a guided discussion session. This was decided taking the social constructivist paradigm to understand and interpret how knowledge is created and transformed by groups of people. This paradigm according to Sandoval (2002) assumes that knowledge is a shared creation between the research participants and the researcher. It also states that researcher's subjectivity and interpretation is necessary to understand human phenomena, in this case, being part of the PA programme as facilitator. The fact that each teacher in the group was engaged both in videoing the microteaching episodes and being videoed allowed the empathetic feelings necessary for creating a challenging but protected learning environment to be developed (Harford & MacRuairc, 2008), where new meanings could be explored and negotiated.

The implementation of the PA intervention was a form of action research, in which the process of interaction is endowed with meaning through what participants do in their actions (Lebak & Tinsley, 2010; Sandoval, 2002; Tabachnick & Zeichner, 1999).

The approach in this research was mixed, based on multi-methods including qualitative and quantitative data collection and analysis techniques. Qualitative techniques were used to gather, analyse and interpret teachers' implicit theories, conceptions and perceptions. These techniques are recommended to explore phenomena seen from the participants' perspective and allow the reconstruction of psychological processes in a retrospective view (Krause, Cornejo, & Radovic, 1997; Krause et al., 2007). Quantitative techniques were used to explore the possible impact of PA on teachers' practice, in order to appreciate differences in participants' teaching before and after PA. According to Jindal-Snape and Topping (2010), that is possible if the researcher identifies observation categories before starting the observation. These categories were organized in an instrument named "rubric to assess quality of explanations" (in appendix 8.8). Likewise, the multi-method approach combined microanalysis of tasks, content analysis, comparative analysis, positioning analysis and a detailed investigation of teachers' performance when they were explaining using quantitative analysis.

3.3. Research design

It has been said that in research designs that are construct-centred such as this study, the intervention programmes should be iterative, process-oriented, involve design products that work in real contexts and need appropriate methodologies (Shin, Stevens, & Krajcik, 2010). This study was centered on the construct of science teacher explanation as an object of development. Considering this idea, this study had a quasi-experimental design. A ten-session PA programme was carried out in three Chilean universities, each one having a comparison and an experimental group. The universities were identified as University 1, University 2 and University 3 (U1, U2 and U3). The researcher played the role of facilitator in PA intervention only, intending to do not influence student teachers' conceptions with the researchers' ones.

Teachers' thoughts were obtained through a PA questionnaire and group discussions. Specifically, teachers' conceptions about the quality of explanations were measured at the beginning and at the end of PA intervention in experimental and control groups. Implicit theories were compared between experimental groups according to their science knowledge.

The quality of the student teachers' explanations was measured in pre and post video-recorded microteaching episodes, using observational analysis and a rubric for assessment. Comparison was made within the experimental groups and between them. A final measurement was taken in some cases in a follow-up study six months after the end of the PA intervention.

In summary, the evidence of the learning process occurrence were sets of data collected from different sources, using different techniques and taking different perspectives. According to Shin et al. (2010), it is important to be clear about what tasks or situations should elicit each evidence of learning. These are presented in the subsequent pages and summarised in page 105 at the end of the chapter in Figure 2.

3.3.1. Peer Assessment intervention design

PA intervention had three main parts, each of them with different objectives and specific task during the sessions. Figure 1 presents a summary of the PA intervention design.

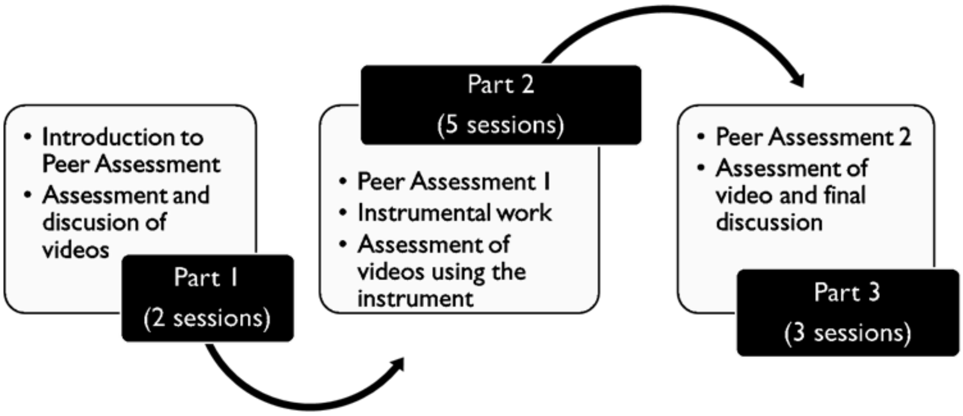


Figure 1: Summary of Peer Assessment intervention design

Part 1: It was designed to introduce the PA methodology and teach to pre-service teachers how to work with it. PA was applied in this stage to video cases in order to avoid possible anxiety in teachers caused by being exposed to and evaluated by a classmate. These videos were obtained from the Chilean Teacher Evaluation System (as detailed in section 3.5.3).

Session 1: The purpose of the programme, the methodology, the participation conditions and the sessions' structure were explained to the student teachers. They signed the informed consent form (in Appendix 8.4) and they watched a ten-minute video in which a young teacher explained the concept of matter. They took notes about the teaching experience to record their observations without any guidance from the facilitator. After, they completed a PA questionnaire (in Appendix 8.5) to evaluate the quality of the explanation and a discussion was carried out. The forms were a primary source of data for this study, following the steps proposed by Kurz and Batarelo (2010). The complete process was repeated with the comparison groups.

Session 2: The student teachers watched a second video and the facilitator conducted a discussion encouraging them to make connections with conceptual or pedagogical issues related to the good and bad practice observed. This part of the intervention was considered useful because another study proved that students who watched a video and discussed the misconceptions with a tutor had larger learning gains than those who watched a video containing just the correct content or the expected performance (Muller, Sharma, & Reimann, 2008).

Part 2: The second part of PA intervention was designed to explore teachers' implicit theories and conceptions about the quality of explanations and to assess their practice to explain.

Sessions 3 and 4: Pre-service teachers were required to individually develop an episode of science microteaching to be presented to their peers in small groups (2-5 participants) avoiding close friends or enemies. This was considered likely to facilitate teachers' confidence in performing in front of peers. Small groups facilitate the scaffolding of knowledge construction among peers (Nicol & Boyle, 2003; Webb, Nemer, & Zuniga, 2002). Besides, it has been reported that small groups had a higher degree of critical analysis and questioning than the whole class working as a group, and students in whole class interventions were more distracted by tangential information in a study by Mayo, Sharma, and Muller (2009). After each explanation peers gave feedback about the performance. This formative assessment was recorded to facilitate the further analysis. Then, a discussion with the whole class about PA and peer feedback was guided by the facilitator. After this 4th session a Sessions assessment form was applied (Appendix 8.6) as a process evaluation to improve implementation aspects.

Session 5: The student teachers watched two selected videos showing one explanation better than the other. They assessed the explanations using the PA questionnaire and were encouraged to establish a link between the practice observed and principles of science teaching, in order to create assessment criteria. The facilitator explored their implicit theories about quality of explanations with this procedure and a rudimentary instrument was created by the participants.

Session 6 and 7: The participants applied the rudimentary instrument to their own recorded microteaching episodes, and they modified or restructured the criteria in a more evaluative language. In cases where they did not want to analyse their own teaching, videos from participants in the other university groups -with their prior authorization- were used.

Part 3: The third part of PA intervention was designed to re-conceptualize the initial participants' conceptions and help them to recognize learning and/or skills development.

Sessions 8 and 9: The same process as in sessions 3 and 4 was done. The student teachers created and performed a second microteaching episode explaining a scientific concept and received peer feedback.

Session 10: The student teachers from the experimental and control groups assessed the same video than at the beginning using the PA questionnaire, following the suggestions of Ferguson (2008) to obtain a post measurement of their conceptions about the explanations. Then, a focus group (in Appendix 8.7) was conducted with the experimental groups addressing their learning or difficulties faced during PA and the factors they considered useful and which ones had an impact in their thoughts and practice.

3.3.2. Research questions

The research questions that oriented this study were the following:

(1) Are the conceptions about the quality of explanations of pre-service science teachers exposed to PA different from those of teachers who have not been exposed to PA?

(2) Are the implicit theories about the quality of explanations of pre-service teachers different according to their science knowledge?

(3) What are the explanation elements that pre-service science teachers use to explain scientific concepts? Are these elements equally modifiable when using PA?

(4) Is it possible to change conceptions and practices about explaining scientific concepts using PA? If yes:

(5) What are the main elements associated with the change process?

(6) Do the changes in explanation practice sustain over the time after PA?

(7) What are the factors (facilitators and obstacles) affecting the transference of good practices to explain scientific concepts into real teaching?

3.3.3. Sampling and participants

In this research the sampling was purposeful and typical cases were selected in sites and participants. However, the universities and individual participants who agreed to participate did so voluntarily, then clearly it is possible that the sample was somewhat biased. The sampling sought to select information-rich cases for study in depth (Patton, 2001) but also it attempted to illustrate what was typical in terms of initial science teacher education programmes and student teachers.

The **universities selected** were typical of Chilean universities which train science teachers. The selection was stratified, looking for the maximum variation possible in the quantity of compulsory science courses given to their students. The selection followed these criteria: (a) That the university was accredited at the moment of contact by the National Accreditation Committee. This is a public organization oriented to verification and promotion of quality in Universities, the careers and programs they offer. This is supported by Chilean law to assure Higher Education quality (Gobierno de Chile, 2011a). (b) That the university had students in a score range 500-550 points in the Chilean entry qualification for university named Prueba de Selección Universitaria (University Selection Test, PSU). This is the average level for student teacher population and the minimum recommended by the Education Ministry for student acceptance. (c) That the university qualification entitles teachers to work in science teaching in middle school (teaching pupils from 9 to 14 years old in the context of Chile).

The **participants** were 38 pre-service teachers, 20 in experimental groups and 18 in control groups. They shared characteristics such as: (a) They chose the science area and were in their final year of training. (b) They had similar practical teaching experience (from zero up to a few weeks). It was important to select people who were in the same course and had a similar level of prior experience teaching, because it has been reported that grade level and the quantity of teaching experience are variables that have an impact on teachers' beliefs (J. Beck, Czerniak, & Lumpe, 2000; Isikoglu et al., 2009) (c) Teachers were 25 years old in average (min.23, max.28, SD 1.7). (d) They represented low and lower-medium socioeconomic status. This is similar to the average level of student teachers in the country. (e) They came from an urban zone of Santiago. This is the capital of Chile and it is where most of the new teachers' training is concentrated.

Even though the percentage of male (40%) and female (60%) was similar in experimental and control groups in the PA intervention, it was slightly different from the national average in primary school science which is 28% and 71% respectively, but more similar to the average in high school science; 38.6% and 61.4%.

The participants were typical of pre-service science teachers in training in Chilean universities. Indeed, as the total of pre-service science teachers in all universities was around 450 in the year of data collection (Palma, 2012), the study sample represents nearly 10% of the population.

The PA intervention was applied as part of the elective courses at the universities and the participation was voluntary. In each university about half of the student teachers agreed to participate. If some of them refused to participate or wanted to quit the programme, they were allowed to without consequences in their marks or study plan. Also, they could participate in PA programme but not be considered in the analysis. However, all participants were interested in taking part in the data analysis. Investigating the possibility of bias in participant self-selection, a detailed comparison was made at pre-test between comparison and control groups' conceptions about the quality of explanations. As mentioned, they assessed a recorded explanation of the concept of matter by a young teacher using the PA questionnaire. After this base-line measurement, they choose whether or not to participate in the intervention. As this initial measurement showed no substantial difference between the participants (details in results section), student teachers that decided not to participate were taken as control group. Their reasons were time constraints or having other elective workshops.

In the follow-up study the participants were selected to give variation in the patterns of evolution after PA: no advance, medium advance and high advance, selected from the group of participant teachers who were working at the moment or in placement. Also, a minimum of two teachers per university were chosen to facilitate the comparison between and within the groups. Nonetheless, in one of the cases (U2) the school authorities did not allow to record the lesson given, but the case was considered equally to participate in the individual interview to assure representativeness. An interesting point to note here is the higher representation of male gender in the followed-up teachers (66.6%) than female (33.3%).

This event might be explained because in the context of Chile the culture is strongly male-centered, then it is easier for men to find their first job than for woman. Actually from the whole group of participants in PA intervention, all the male teachers were working after 6 months of their graduation, although around 50% of women only were working and just 40% of them were in educational jobs. Then, it was not surprising to have a different gender distribution, as summarised in Table 5.

Table 5: Sample's characteristics experimental group

Group	Number of participants	% Female	% Male	Age (average)	Age (standard deviation)	Age (min)	Age (max)
PA intervention	20	60	40	24.9	1.7	23	28
Follow-up	6	33.3	66.6	24.8	1.83	23	28

3.4. Measurements, data collection techniques and instruments

3.4.1. Research variables and their measurements

The study variables and their measurements were:

(a) The teachers' conceptions: They were understood as the explicit beliefs teachers hold about conceptual explanations. The researcher accessed them through PA questionnaire (see details in Section 3.4.3) and feedback session discussions. The PA questionnaire was applied to the experimental and control groups before (pre-test) and at the end of PA (post-test). The change in teachers' conceptions was assumed when their evaluative comments in the post-test were qualitatively different compared with the pre-test. This was a qualitative variable.

(b) The teachers' implicit theories: They were the implicit principles underpinning the teachers' conceptions. As they were implicit, the researcher accessed them through analysing the feedback sessions and the sessions to construct the assessment rubrics that each group used to assess their peers' performance. This was a qualitative variable.

(c) The quality of teachers' explanations: It was measured with a rubric (see details in Section 3.4.4) and it represents the extent to which teachers' explanation developed in the microteaching episode fulfilled the quality criteria purposed. The assessment was applied in the experimental group at the beginning and at the end of PA. In some cases it was also applied six months after PA in the follow-up study. This was taken as a quantitative variable.

(d) The teachers' science knowledge: It was determined by the number of science courses that universities offered to student teachers to enable them to work as science teachers. This was considered as a categorical quantitative variable. University 1 gave fourteen (high science knowledge), university 2 gave nine (medium science knowledge) and university 3 offered four (low science knowledge).

To ensure validity in the data gathered, there were different instruments and sources to collect them yielding a wider view of the phenomena. These instruments also helped to evaluate the intervention and its results. They are described in the following pages, mentioning also the moment in which they were applied.

3.4.2. Data collection techniques

(a) Participant observation: Observation is a method of data collection that employs the sense of vision as its main source (Jindal-Snape & Topping, 2010). It could be categorised as naïve or scientific depending on its orientation by an objective. It could be participant or not participant (according to the degree of researcher involvement in the phenomena), structured or unstructured (depending on how the aspects to be observed are guided), natural or laboratory (if the researcher creates or does not create the observation conditions), among others (Sarantakos, 2005). In the case of this research, semi-structured and structured participant observation was used to have an accurate view about the PA feedback sessions and the teachers' performance during all the intervention - in verbal and non-verbal communication - and to guide the discussions. Participant observation accesses to the on live contact with the reality, phenomena or research interest, and the researcher is part of the context where the phenomena is happening (Sandoval, 2002).

(b) Product or artefact analysis: It is usually defined in the social sciences field as the analysis of the visible expressions of a culture, behaviour patterns, concrete productions or mental representations (Gagliardi, 1990). The products analysed in this research were the teachers' notes they took during the video observing and the rubric created from their personal point of view about quality criteria to explain scientific concepts.

(c) Focus group: It is a semi-structured discussion focused on a few topics in a small group of relevant people for the research objectives -usually between six and eight- (Sandoval, 2002). In this research it was applied at the end of PA intervention as a final post-evaluation, designed to explore teachers' perceptions about the intervention and recognitions about learning they could have had during the sessions. It was guided by the facilitator and it was audio recorded to make the data analysis easier (see the question guide in Appendix 8.7).

(d) Individual interviews: According to Taylor and Bogdan (1993), the individual interview is a research tool in which the participant and researcher meet to discuss deeply about a particular topic. It could vary from a non-structured flow to a totally structured one. In this study it was done to explore teachers' views about the transition from initial teacher training to real teaching experience and a semi-structured guide was chosen. The interviews had a set of questions according to the aim, but they differed in form, order and extension depending on the interaction with the participant (question guide in Appendix 8.7).

3.4.3. Data gathering instruments

(a) Peer assessment questionnaire: It was a device that helped to analyse the conceptions that student teachers held about quality of conceptual explanations. It was applied at the first, second and third parts of PA. It included a six-level semantic differential scale (from "very bad" to "very good") to allow teachers assessing the quality of peers' explanations, and open questions to justify the assessment (see the form in Appendix 8.4). To ensure the instrument was satisfactory, a preliminary version was tested in a pilot study and through expert panel submission, and any necessary corrections were made. At the beginning of the intervention it was applied in order to familiarise pre-service teachers with the instrument and obtain a measurement of their conceptions before PA started (in this stage it was applied to videos). In the second part of PA, the questionnaire was used to help teachers when assessing the peers' conceptual explanations. At the end of PA it was used to access to teachers' conceptions, using the same video than at the beginning of the project.

(b) Sessions assessment form: Every four sessions a formative evaluation was carried out to evaluate the implementation process of PA from the participants' perspective. The purpose was to get feedback through asking if they perceived the sessions' objectives were reached. This allowed the researcher determining whether the intervention was running in the expected way or not. It included closed questions in a Likert Scale and an open space to write any other aspect not contemplated before (see the instrument in Appendix 8.6).

(c) Rubric to assess the quality of explanations: The rubric was an instrument created to assess the participants' explanations. The ten indicators were formulated based on the information gained from the literature review and teaching frameworks (see details of rationale and reference to authors and researchers in the next Section 2.6.7). A preliminary version was discussed with expert science teachers from the United Kingdom and Chile, then tested in a pilot study and further sent to an expert panel to review it before the intervention. Any necessary corrections were made to increase the instrument's reliability and validity. The final version (in Appendix 8.8) contained three levels of achievement following the structure suggested by Chilean Ministry of Education to assess initial teacher education standards (Gobierno de Chile, 2001). The rubric gave a score from zero to 20. Scores between 0-6 were taken as low performance, 7-13 as medium and 14-20 as high. It was applied to all microteaching episodes and participants' lessons in the follow-up study.

3.5. Procedure

3.5.1. Negotiation with universities

The researcher visited the authorities from the universities involved six months in advance in order to explain them the project, to get the permissions to run the PA intervention as a practical seminar and the necessary information (i.e. career programmes, access to students' previous marks in science courses, etc.). To obtain these meetings, an introduction letter was sent to the Deans of Education (see Appendix 8.1) and a PowerPoint™ (Microsoft, 2010) presentation was generated for this proposition (in Appendix 8.2). Three permits were obtained for the researcher to work with pre-service science teachers. Initially, it was agreed with the universities to run the PA within a regular course called "science didactics", working in collaboration with the teacher in charge and randomly assigning the student teachers to the control and experimental groups. However, a month before the implementation of the PA programme there was a significant higher education student demonstration and strikes that did not permit conducting the normal university lessons as planned. Most of the universities stopped their activities and consequently, the authorities of the selected universities preferred to run this project voluntarily for the students. This was contrary to the initial agreement, but it was understandable that in this context they gave priority to completing the students' compulsory curricular activities and any extra-curricular activity had to have a voluntary character.

3.5.2. Ethical considerations

This study was conducted according to the "Research Ethics: Code of Practice" from University of Dundee (University of Dundee, 2007). The ethics committee reviewed and approved the research layout, indicating they had no concerns about the ethical implications of the study for the participants.

The participants were voluntary and they received an invitation to take part of the study with a full explanation about the research features and the tasks they were asked to fulfil. This was given orally by the researcher but also written in a participant information sheet (in Appendix 8.3). The student teachers signed an informed consent form to give their authorisation to the data gathering (in Appendix 8.4). If some students wanted to quit the programme they were allowed to do so without consequences in their marks or study plan.

Also, they could participate in the programme but not be considered in the analysis if they wanted. Even so, all the participants expressed their motivation to be incorporated in the data analysis.

All the data were marked with a student identification number that was allocated by the researcher in order to match the different products and documents with the participant but keeping their identity anonymous. The video-audio material was recorded and all data records were saved on a password protected computer network. A back up was kept in a secure office in University of Dundee, School of Education, Social Work and Community Education to avoid accidental loss or damage. After the research was finished and the results were generated and reported, the data records were destroyed. All data was treated confidentially, the only ones having access to them were the researcher, the two research assistants and supervisors.

3.5.3. Equipment and materials

To run this research and the intervention, it was necessary to have per each university a classroom with good acoustics, one data projector, one laptop, one audio recorder, two camcorders, two stands and battery replacements. It was also necessary to have printing and photocopying facilities to ensure enough copies of the informed consents, information sheets and PA questionnaires.

During the PA intervention the microteaching episodes were recorded with two video cameras (one for backup), while in the follow-up study the lessons were recorded from the back of the classroom with a laptop only. This was decided following advices from Geelan (2009), to reduce the impact on teachers and pupils and make the recording as unobtrusive as possible.

Finally, it was necessary to have videotaped science real lessons and one videotaped lesson specially designed to train teachers in PA showing different quality of the explanations. These videos were facilitated by the National Teacher Evaluation System in Chile, because this work team had a large number of videos they often use for the validation of their instruments. The videos were from in-service teachers who were not included in the evaluation process and manifested their will to give access to record one of their lessons. These teachers signed an informed consent authorising the use of their videotaped lesson for research purposes, teaching purposes and for developing assessment instruments in pilot studies. Although the National Teacher Evaluation System had several videos of all teaching subjects, for the pilot study of the present research videos of science teachers aged between 25 and 30 years old were selected only. This was to assure the student teachers of this study would recognise the teachers in the videos as peers in terms of teaching experience. The videos were used first in the pilot study of the rubric and then some of them were edited to be part of PA intervention.

3.5.4. Triangulation

Triangulation is a technique intended to increase the strength in research rigor. Besides, it is recommended to enhance the trustworthiness of analysis by providing a more inclusive and complete view (Denzin & Lincoln, 1994). To ensure the results were not taken from a partial view or they were mainly subjective, in this research triangulation was accomplished in three of the four types proposed by Patton (2001):

(1) Methods triangulation: The researcher used a variety of sources and techniques to collect data. It has been said that this minimises the risk of bias and limitations of any individual method by compensating with the strengths of another method (Denzin & Lincoln, 1994).

(2) Data triangulation: The researcher relied with multiple kinds of data. This was recommended by Conner (2010) when there is information in multiple directions, to allow data sources' converging and identifying different ways the phenomenon was expressed.

(3) Researcher triangulation: The instruments that contained interpretative data (such as videotaped microteaching episodes, feedback sessions) were triangulated by researcher. A 5% of all qualitative data sets were randomly selected and checked by a second researcher, in a blind process. If the results varied between them, they were discussed until consensus was reached. To apply the rubric to the participants' explanations, 100% of the videos followed the triangulation process due to their relevance for this research. The agreed scores, assignations and marks were used for the data analysis and to report results.

3.5.5. Measurements timetable

A summary of the time when the measurements were taken is presented in Table 6:

Table 6: Intervention and measurements time table

Activity/measurement	Sep '11	Oct '11	Nov '11	Dec '11	Jan '12	Feb '12	Mar '12	Apr '12	May '12	Jun '12
Intervention experimental group U1	X	X	X							
Measurements U1	X	X	X						X	
Control group measurements U1	X		X							
Intervention experimental groups U2 and U3		X	X	X						
Measurements U2 and U3		X	X	X						X
Control groups Measurements U2 and U3		X		X						

3.6. Data analysis techniques

The data analysis process was carried out according to the steps proposed by LeCompte (2000). The first step was to tidy the data up and to prepare it for coding. It meant making copies of all data, putting all field notes, observations and documents into a file in a certain order according to participant number and dates of creation. The second step was to find items: the specific pieces of information in the data set that the researcher coded, counted and assembled into research results. Items were identified because they were frequent (frequency criteria), because they never appeared even though researcher might think it reasonable that they would (omission criteria), or because the participants told the researcher the items existed (declaration criteria).

In general terms after the item identification, the researcher organized the items into groups or categories by comparing, contrasting, mixing and matching them. The purpose of these activities was to clump together items that were similar (similarity) or appeared together (co-occurrence), or in patterns whose existence was confirmed by other pieces of data or information (triangulation). Once patterns were identified, the next step was determining taxonomies that could be collated together in meaningful ways and assembling structures or linking patterns that built an overall description of the phenomena being studied.

3.6.1. Analysis of conceptual explanations

As it was mentioned, the explanations were video recorded. This decision was taken following the suggestions of Sevian and Gonsalves (2008) considering that body language and other media such as images, diagrams or others used during the lesson were important for the analysis as proved earlier by Geelan (2009). To encourage the reliability of the observational data analysed 100% of the videos were analysed for the researcher as mentioned before. Another observer double marked all the explanations present in the videos. It has been said that 70 or 80% of agreement is satisfactory between observers (Jindal-Snape & Topping, 2010). In this study the inter-judge agreement was 80% in average. However, all the divergent aspects were discussed after marking to reach consensus on the final mark. The conceptual explanation as a teaching device was analysed with the statistical analysis tools because it was measured in quantifiable criteria by the rubric.

3.6.2. Narrative analysis

According to Carter and Doyle (1996), a narrative can be considered to be a methodological device that can interpret experiences or events that reflect a more general understanding. The narrative analysis can be applied to classroom contexts considering they are composed of different events and experiences for the teachers and the pupils (Rodrigues, 2010). Likewise, it is important to mention that narrative enquiry is not simply an account about what happened in special diverse moments or circumstances. But the focus of narrative enquiry is on how people make sense of what happened (France, 2010) and for this reason the approach was considered in this research, taking into account that to understand teachers' thoughts it is necessary to deconstruct how they constructed their meanings.

As stated by France (2010), a successful narrative should allow the reader to make their own connections by way of an interpretation with their own story, taking as first source the language, discourse or talk somebody is using in determined moment. The literature related to this theme shows that language and talk in science classrooms can be studied and typified in different ways. Language can be considered in terms of where it happens, who is responsible for generating it and how it is done, for instance, involving teacher and/or pupil engagement and written and/or spoken language. Any kind of explanation should be understood in the general context of interpersonal communication, because in teacher explanation communicative strategies are involved (Faye, 2009). Pursuant to this author, if explanations are analysed as an intentional act of communication, they could be seen as context-bound, directed, intentional, potentially persuasive and determined by the public rules of speech.

The positioning analysis in the context of narratives was besides applied in this research. A position is a metaphorical term that describes a psychological location or space which a person occupies in a workplace or other conversation (Phillips & Hayes, 2008). In accordance with Redman & Fawns (2010), the position is expressed in language or gesture in a specific moment or different moments. The position is likely to influence what people subsequently say and do. Underpinning positioning theory there is the idea of socio-psychological tools mediating in social factors, highlighting how the collective representations may be actively interpreted by the people in a moment (Daniels, 2004). This analysis has been used with teachers in primary science classrooms to identify the relationship the speaker had with the topic (Redman & Fawns, 2010), and also in people's talk, observing what they pointed to, draw or align themselves with or against the shared values (Moghaddam, Harre, & Lee, 2008). This analysis was used in this research in feedback sessions and data from student teachers discussions.

3.6.3. Statistical analysis

The numeric data gathered from PA questionnaires and the rubric were analysed statistically. This was done to appreciate whether differences existed when comparing pre-post intervention times in experimental and comparison groups, and between experimental groups including the follow-up.

Specifically, statistical descriptive and inferential analyses were run. The analysis of variance (ANOVA) is regularly used to compare groups in a particular variable which is partitioned into components attributable to different sources of variation (Field, 2005). In this research, one-way ANOVA was used to compare the quality of explanation in experimental groups in University 1, 2 and 3 according to their science knowledge and the quality of explanation in experimental groups (grouped together) in pre and post microteaching episodes. The Statistical Package for the Social Sciences (IBM Corp., 2010) software version 19 was used.

Frequency analysis was used to describe statistically control and experimental groups according to the pre and post PA questionnaire scores (semantic differential scale). Also, descriptive analysis was done to find out if the rubric criteria were discriminatory in the instrument validation process. In this process the rubric internal coherence was also investigated using Cronbach's Alpha indicator.

In the case of process evaluation, frequencies were used to take decisions in the programme implementation. Finally, descriptive analysis was applied to count the appearance of categories in evaluative comments given by pre-service teachers during feedback sessions. To obtain this categorization, qualitative analysis were conducted to organise the data. These are described in the following pages.

3.6.4. Analysis of pre-service teachers' thoughts

The pre-service teachers' notes during PA sessions, the open questions in PA questionnaires, the student teachers' evaluative comments during discussions and feedback sessions, interviews and the final focus group were analysed using two different techniques, in order to perceive possible differences in the participants' thoughts: constant comparison analysis developed by Glasser & Strauss (1967) and content analysis described by Neuendorf (2001).

3.6.4.1. Constant comparative analysis

The constant comparative analysis leading to Grounded Theory was developed by Glaser & Strauss (1967) as a methodological approach that follows a cyclical process of induction, deduction and verification of data through a set of strategies. According to Valanides (2010) it is generally used to dissect, conceptualize and categorise qualitative data. Patterns in data are revealed and constantly refined in the categorization process, following these stages: comparing incidents applicable to each category, integrating categories and their properties, delimiting the theory and writing the theory (Dye, Schatz, Rosenberg, & Coleman, 2000). Valanides (2010) mentioned that the whole process implies attempts to identify incidents, code them accordingly, and organize them into categories using a never-ending comparison. Also, the names of categories usually change because the properties or characteristics of a category are progressively identified and developed based on data, according to different new categories or subcategories that are identified.

The processes of coding are named open coding, axial coding and selective coding, but the division between them is artificial because the types of coding do not take place in a compulsory sequence (Glaser & Strauss, 1967). All these processes were considered in this research. Open coding is usually used in the first level of abstraction, where everything in the data is coded to find the first categories or patterns. It implies naming and categorizing the phenomena by closely examining data and incidents, in order to identify similarities and differences (Valanides, 2010). Axial coding is an inductive and deductive process guided for constant comparisons within the data. The purpose is to discover and relate categories in terms of a paradigm model (Strauss & Corbin, 1990). Selective coding constitutes the final integration or the final leap between creating a list of concepts and producing a theory for explaining a phenomenon (Valanides, 2010).

3.6.4.2. Content analysis

Content analysis is a technique of text analysis, not only considering written texts but also those painted, recorded, spoken, etc. (Andreu, 2001). It has been defined from 1952 in many ways, focusing in different aspects. However, there is an integrative definition that established content analysis as the conjunction of analysis techniques of communication that tends to get indicators through systematic and objective procedures of content description of the messages. It allows inferring knowledge related to the conditions or the context in which these messages were produced (Bardin, 1996).

Andreu (2001) highlighted one of the most important utilities of the content analysis for this research; the possibility to perceive hidden meanings or indirect messages in the text. He indicated these inferences are referred fundamentally to the symbolic communication or data meaning that are in general different from the observable data. According to his view, to infer is to deduct what is inside a text, extracting conclusions and explanations that are implicitly contained, looking for the components, the internal links in the information and its transformation.

Content analysis in the present research started by the researcher with a thematic analysis of the data collected from the sessions to create student teachers' rubric. Thematic analysis is considered by Braun & Clarke (2006) as a poorly demarcated but widely used qualitative analytic method. They defined it as a method for identifying, analysing and reporting patterns or themes within data, through the organization and description of the data set in detail. Thematic analysis can be seen as a foundational method for qualitative analysis or it can be considered as a method in its own right, although it has also been considered as a non-specific method but as a tool within different analytic traditions such as the Grounded Theory (Boyatzis, 1998; Ryan & Bernard, 2000). In this research thematic analysis was considered as the first stage of a major analysis method in PA questionnaires, focus groups and interviews data, but in feedback sessions' data it was chosen as the main analysis method used, followed by positioning analysis (Redman & Fawns, 2010).

One of the advantages of thematic analysis is its flexibility. It allows the researcher to determine themes and prevalence in a number of ways. Although the explicit level of analysis is more usual, it could be used at a latent level (Patton, 2001). This was used in this research, identifying the themes underpinning the data surface, describing them and interpreting them, attempting to theorise the significance of the patterns and broader meanings and implications. The thematic analysis steps are also flexible and involve a constant moving back and forward within the entire data set. In this research the steps used were: familiarizing the researcher with the data, generating initial codes, searching for themes, defining, reviewing and naming themes, as suggested by Braun and Clarke (2006).

After the thematic analysis, the second stage in the content analysis was the elaboration of indicators in the information and the definition of analysis units. They were the meaning nucleuses used for classification and counting. The rules of numeration or counting were stated following the ones suggested by Porta & Silva (2003); presence or absence of a determined code, frequency of occurrence of a code, intensity, contingency, occurrence order, density in a determined text (percentage of the total of frequencies that a code had in a text) and concentration (percentage of the quantity of different codes found in that text). At the moment of counting, the emergent guides from the relations that linked the different categories and indicators were determined. After that, an inventory of meaning units was constructed, that was a type of radiography of the present ideas in the material reflecting the categories, conceptions and implicit theories from the participants. This radiography was the fundamental basis to explain and describe the data obtained in the results (see an example in Table 11, page 116). It is important to mention that the data obtained from the analysis methods mentioned were imported into a qualitative data analysis software package called NVivo (QSR, 2011). It provided a data management tool including different levels of analysis (Denley & Bishop, 2007). It simplified the management of large amounts of data coming from different sources.

Finally, the findings obtained from different sources of data about the same object of analysis were integrated in a network in a process of theoretical consolidation, following the recommendations of the content analysis of networks model (Hoey, 1991).

A summary of the research questions, objectives, techniques, instruments used and analysis method are presented in Figure 2 in the next page, to illustrate the connections between these elements.

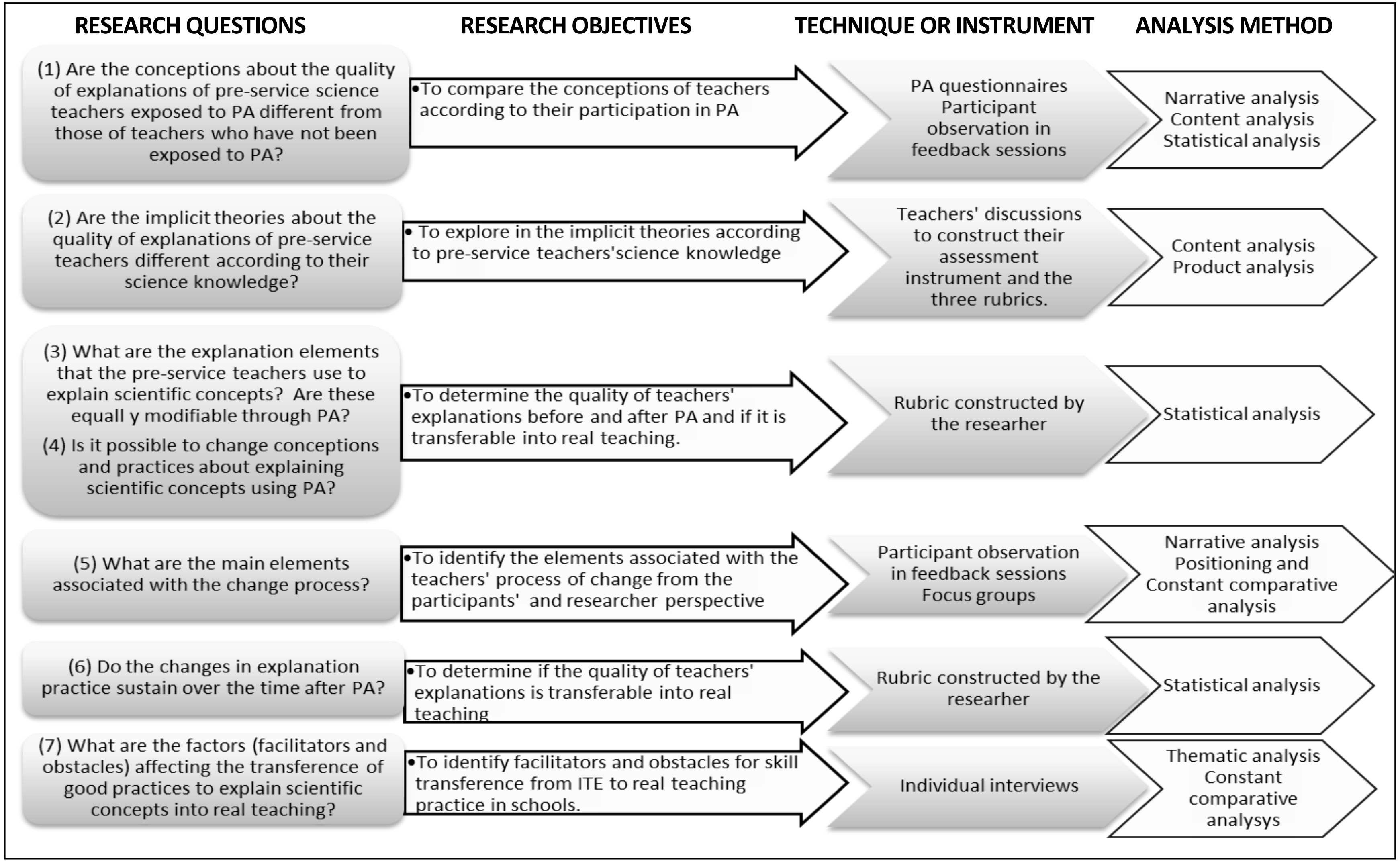


Figure 2: Summary of research questions, objectives, techniques or instruments and analysis methods

4. Results

This chapter presents the main results obtained in the research, organized into three studies. The first study shows the construction process of the rubric used to assess the quality of teacher explanations, including its validation and descriptive results. The second study describes the peer assessment (PA) intervention findings, comprising the conceptions that the participants held about the quality of the explanations, how these conceptions changed towards the development of pedagogical content knowledge (PCK) and the elements that promoted the changes in student teachers' theories and practices. Finally, the third study presents the results of the follow-up of participant teachers' explanations in real classroom context (called beginning teachers in this moment), the facilitators and obstacles for the skill transference.

4.1. Study 1: Construction of an instrument to assess the quality of teacher explanations

4.1.1. Construction and pilot study for validation

As mentioned in the previous chapter, section 3.4.4, a rubric was created from the research literature in science education and general teaching frameworks to assess teachers' explanations. This instrument allowed translation of observational evaluation into quantitative data.

In the first draft of this instrument, seventeen criteria were found possible to be applied to teacher explanations: (1) indicating the explanation objectives; (2) explanation introduction or conceptualization (3) clarity; (4) coherence and cohesion; (5) sequence; (6) accuracy; (7) connection with pupils' prior knowledge; (8) connection with pupils' daily life; (9) usage of examples; (10) usage of experiments or demonstrations; (11) usage of graphs or images; (12) usage of metaphors or analogies; (13) usage of simulations or models; (14) usage of body gestures; (15) inflections in voice; (16) usage of pupils' mistakes or common mistakes as a learning opportunity; (17) summary of the explanation .

These criteria were adjusted to operate in three levels of achievement (not achieved, half achieved and achieved). Each level contained quality elements that might be present and observable during teaching. Zero points were given to each level not achieved, because the explanation did not present the elements or the elements were poor quality. One point was assigned to the half achievement level. It implied the explanation presented the criterion but not with the expected quality. Two points were marked for the full achieved level. It meant the explanation fulfilled the criteria with quality elements as expected.

The seventeen-criterion version was discussed in terms of content with two expert teachers in science education from Chile and the United Kingdom before being applied. Corrections were made and some criteria were merged following the experts' suggestions. The key to merging the criteria was the function they might have in an explanation, i.e. the examples, images, graphs, experiments or demonstrations, which at the beginning were individual criteria, were grouped together because they may illustrate or clarify certain aspects or properties of the concept being explained. The same process operated with the criteria 7 and 8, 12 and 13, 14 and 15. Otherwise, the initial criteria 1, 2 and 17 had to be removed from the instrument because they would be not possible to observe in a microteaching episode due to its concentrated and simulated character. Thus, a ten-criterion version was obtained.

This second version of the rubric was tested in a pilot study with 17 science teacher explanations videotaped in different real teaching contexts driven by beginning teachers as mentioned in the procedures of this research, chapter 3 (methodology), section 3.5.3. The criteria tested were (1) clarity; (2)coherence and cohesion; (3) sequence; (4)accuracy; (5) sufficiency; (6) connection with pupils' experience; (7) usage of analogies, metaphors, simulations or models; (8) usage of visual representations; (9) usage of non-verbal language and (10) usage of pupils' mistakes or common mistakes as a learning opportunity.

Although the correlation matrix (Table 7) showed that very few of the rubric items (SQ) were significantly correlated (using Pearson's correlations), this was not considered a problem because it was assumed the items were measuring different aspects of the teaching practice to explain scientific concepts. Indeed, finding this result in beginning teachers was to certain extent expected, because they were supposed to have a heterogeneous pattern in the components of their skill of explaining, as they were not expert teachers yet.

Table 7: Correlation Matrix of the rubric (second version)

	SQ1	SQ2	SQ3	SQ4	SQ5	SQ6	SQ7	SQ8	SQ9	SQ10	Corr. total
SQ1		.365	.236	.067	-.209	.457	.460	.309	.240	.127	**.584**
SQ2	.365		-.150	.038	-.262	.259	.181	.299	.136	-.278	.363
SQ3	.236	-.150		-.251	.190	.108	.174	.073	.194	-.116	.165
SQ4	.067	.038	-.251		.403	.203	.179	**.526**	.055	-.136	**.621**
SQ5	-.209	-.262	.190	.403		.137	.071	.197	.246	-.313	.391
SQ6	.457	.259	.108	.203	.137		-.045	.409	.377	.200	**.605**
SQ7	.460	.181	.174	.179	.071	-.045		.142	.149	-.043	**.529**
SQ8	.309	.299	.073	**.526**	.197	.409	.142		.135	-.248	**.698**
SQ9	.240	.136	.194	.055	.246	.377	.149	.135		.105	**.513**
SQ10	.127	-.278	-.116	-.136	-.313	.200	-.043	-.248	.105		-.051

Correlation is significant at the 0.01 level (2-tailed) Correlation is significant at the 0.05 level (2-tailed)

The internal consistency of the instrument was calculated using Cronbach's Alpha (Table 8) to find out how closely related the criteria were. According to this indicator, the rubric was not found to be highly reliable (α = 0.60, n=10). However, as every item contributed to reliability and they were supported by literature review, it was decided to reorganize the criterion elements based on their correlations instead of deleting some of them.

Table 8: Cronbach's Alpha of the rubric (second version)

Cronbach's Alpha	Number of Items
.604	10

The refined rubric (third version) was validated through the views of an expert panel, (the instrument construction team from National Teacher Assessment System in Chile). They suggested language modifications, specifications and re-ordering in some criteria. All these suggestions were effected, giving a fourth version of the instrument.

The fourth (final) version of the rubric (in Appendix 8.8) was used in the PA intervention with twenty pre-service teachers. Microteaching episodes were videoed as pre-post measures, giving a total of forty teacher explanations to be analysed. As mentioned before, to assure reliability of the observational data, 100 % of the videos were analysed by the researcher and at least one of two carefully trained and chosen research assistants. Approximately one third of the forty videos were analysed by the female assistant and the other two thirds with the male assistant. Inter-judge agreement was calculated between the researcher and the observers. With the female observer 81.8% of absolute agreement was reached and with the male observer 78.9%. The average agreement was 80.35%. All the divergent aspects were discussed after marking, until a 100% consensus was reached. The agreed score was used in the statistical analysis in every case.

The internal consistency of the instrument was calculated with forty microteaching explanations using Cronbach's Alpha (Table 9). The rubric was found to be reliable (α = .77, n=10) and all the items contributed to its internal consistency. Any indicator, if deleted, would have substantively increased the reliability of the final version of the rubric.

Table 9: Cronbach's Alpha of the rubric (final version)

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	Number of Items
.77	.78	10

Taking into account the small case number this study had, it was not likely to find significant correlations between the items. However, in this version of the rubric the Pearson's correlations were better than in the previous one (Table 7 and 10 respectively). Around the half of the correlations between the items were statistically significant (p<.05), and all the items were statistically significant correlated with the total score (Table 10).

Table 10: Correlation Matrix of the rubric (final version)

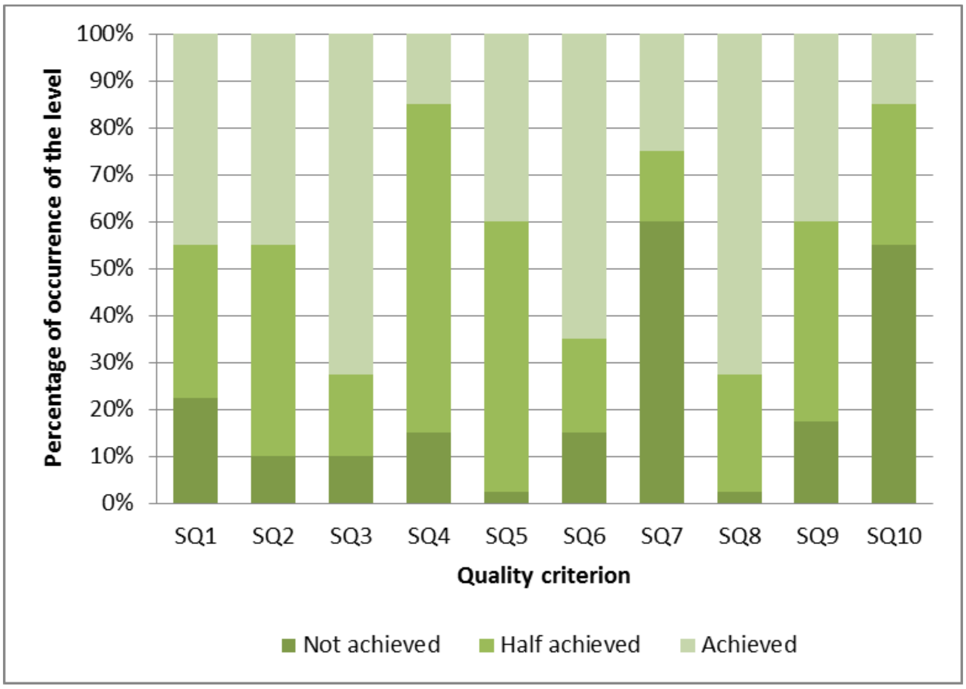
	SQ1	SQ2	SQ3	SQ4	SQ5	SQ6	SQ7	SQ8	SQ9	SQ10	Corr. total
SQ1		.138	-.030	**.462**	.156	**.576**	**.339**	.168	**.348**	**.413**	**.653**
SQ2	.138		**.479**	**.558**	**.627**	**.361**	.175	.165	.309	.291	**.669**
SQ3	-.030	**.479**		.208	**.471**	.281	.078	.112	-.085	.103	**.414**
SQ4	**.462**	**.558**	.208		**.514**	**.431**	.214	**.358**	**.378**	.124	**.693**
SQ5	.156	**.627**	**.471**	**.514**		**.348**	.179	.230	.170	-.064	**.568**
SQ6	**.576**	**.361**	.281	**.431**	**.348**		.237	.132	**.349**	**.413**	**.730**
SQ7	**.339**	.175	.078	.214	.179	.237		.161	**.330**	.136	**.541**
SQ8	.168	.165	.112	**.358**	.230	.132	.161		.183	.147	**.413**
SQ9	**.348**	.309	-.085	**.378**	.170	**.349**	**.330**	.183		.216	**.573**
SQ10	**.413**	.291	.103	.124	-.064	**.413**	.136	.147	.216		**.520**

Correlation is significant at the 0.01 level (2-tailed) Correlation is significant at the 0.05 level (2-tailed)

4.1.2. Rubric pilot study results

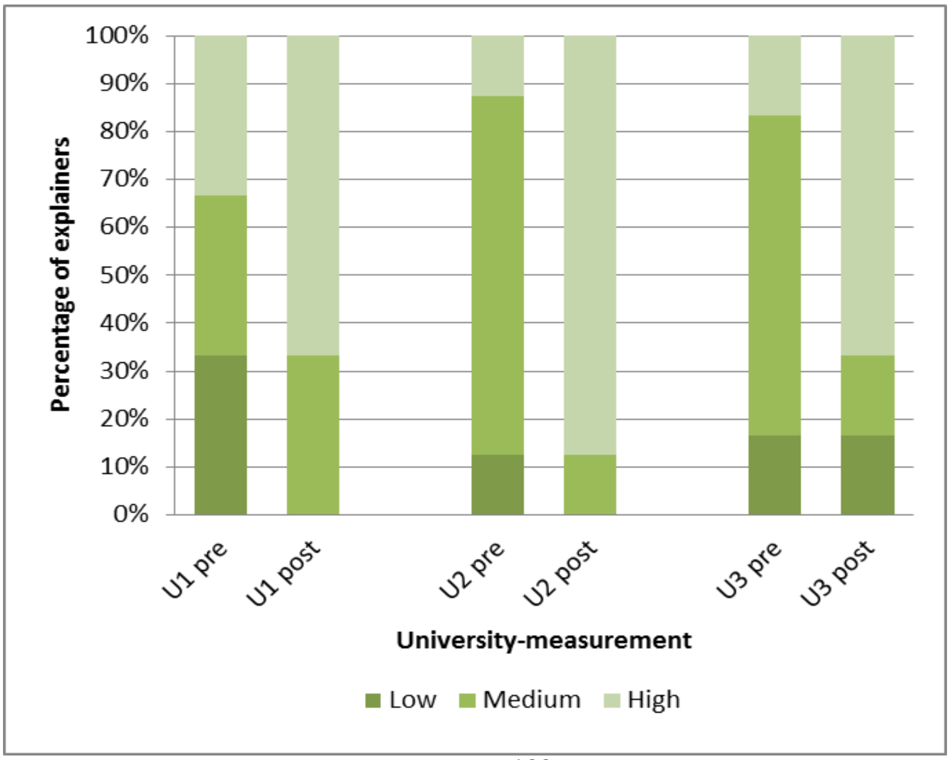
The application of the rubric to the microteaching episodes pre and post PA allowed to test the metric characteristics of the instrument. Figure 3 shows the three levels (not achieved, half achieved, achieved) of each item of the instrument were at least used once. This fact supports their applicability and pertinence to assess different aspects of science teacher explanations. Likewise, all criteria were discriminatory.

Figure 3: Rubric criteria distribution by level of achievement



Besides, it was possible to identify low, medium and high quality explainers within each university. The distribution of student teachers in low, medium and high quality of performance groups showed changes between the pre and post measurements. In the three universities after PA there were more student teachers in the high quality performance group, as shown in Figure 4.

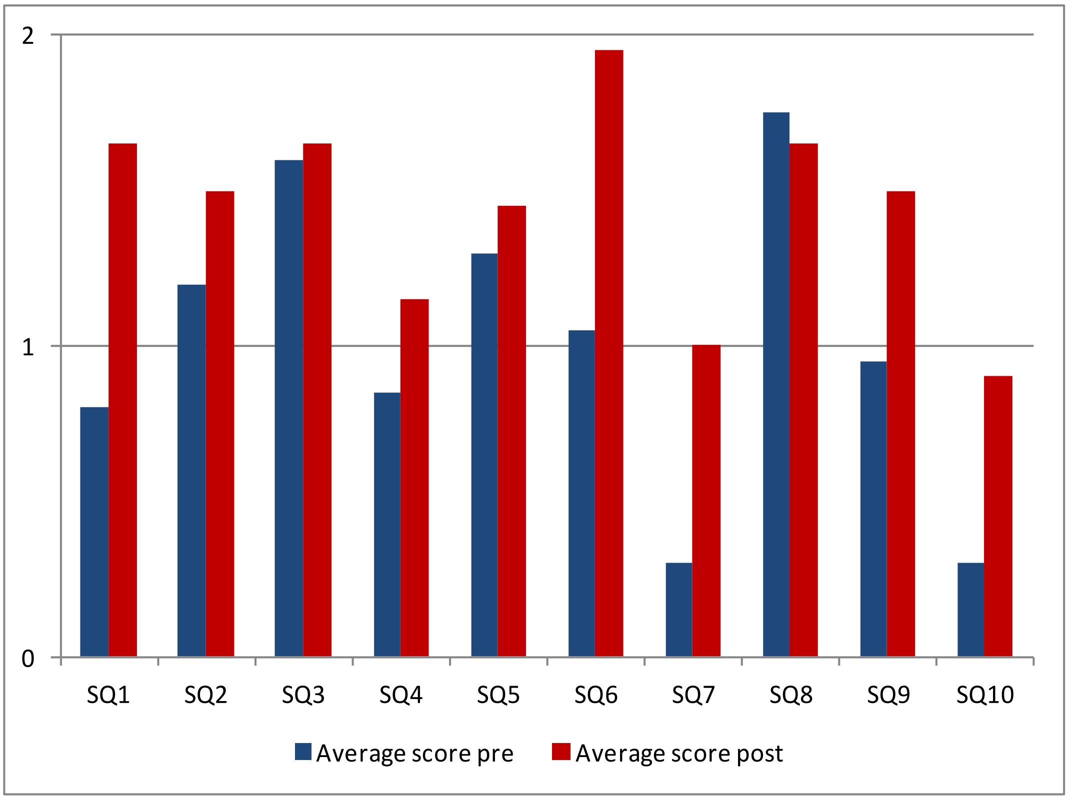
Figure 4: Overall score pre-post by university



Otherwise, a quality of explanation pattern was possible to be established. In general terms, this pattern was lower at the beginning of the intervention and higher at the end of it. Figure 5 shows there were criteria with an important improvement (i.e. SQ1, SQ6), with almost no improvement (i.e. SQ3, SQ8) and a criterion with a decrease (SQ8). The complete criteria description was presented in Chapter Methodology, section 3.4.4.

This suggested the rubric was suitable to identify strengths and weaknesses in the quality of teacher explanations. Moreover, the instrument was sensitive enough to detect changes in teachers' performance to explain scientific concepts.

Figure 5: Average score in rubric pre-post by quality criterion



SQ1	Clarity	**SQ6**	Connection with pupils' experience
SQ2	Coherence and cohesion	**SQ7**	Metaphor, analogy, simulation or model
SQ3	Sequence	**SQ8**	Example, demonstration, graph, image, experiment
SQ4	Accuracy	**SQ9**	Gesture and voice
SQ5	Sufficiency	**SQ10**	Misconception illustration

To summarise, after development, this rubric was a valid instrument to assess different elements of science teacher conceptual explanations, which together might act as a skill of explaining. Therefore, the instrument allowed identifying pre-service teachers' strengths and weaknesses in their performance to explain scientific concepts, which configures the rubric as a tool for skill diagnosis.

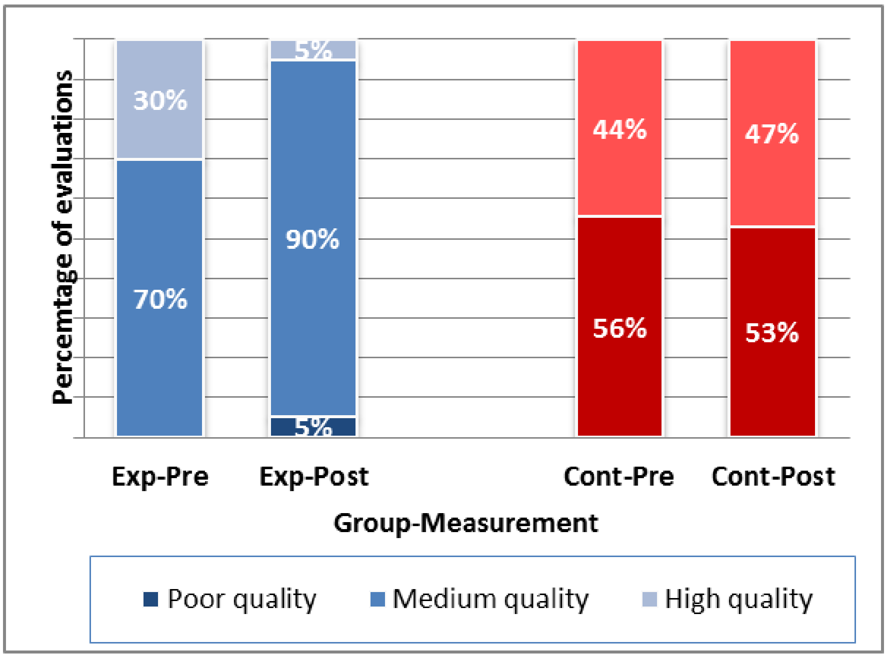
4.2. Study 2: Peer Assessment intervention

To determine whether or not teachers' conceptions about the quality of explanations were modifiable by PA, a comparison between control and experimental groups was done. In the three universities a video recorded explanation was presented at the beginning and at the end of PA to allow teachers analysing it, judging its quality it in a six-level semantic differential scale and justifying the evaluation (Appendix 8.5). As it was mentioned before, the video showed a young teacher explaining a scientific concept (definition of matter).

4.2.1. Student teachers' judgments

Regarding the teachers' judgment, in the six-level semantic differential scale (from "very bad explanation" to "very good explanation") the two first scale levels were interpreted as low quality perceived, the following two as medium and the last two as high. Frequency counting and trend comparison were used to analyse the data. The results obtained indicated that before PA intervention 70% of the experimental group perceived the explanation as medium quality and 30% as high quality. After PA, 25% of this group gave a more critical judgement. In contrast, before PA 56% of the control group considered the explanation as medium quality and 44% as high quality. After the time of intervention the control group's judgement remained almost stable (Figure 6).

Figure 6: Differences in teachers' judgements Pre-Post experimental and control groups



4.2.2. Student teachers' conceptions held

The evaluative comments given by the student teachers in PA questionnaires were coded following the indications of Glaser (1965) regarding the constant comparative method of analysis. The comments were organised in four categories (agreed with another researcher) according to the central aspect the teacher took into account to assign the evaluation.

(1) General Aspects (GA): This category grouped together the aspects mentioned by the participants not related with the way the teacher delivered a conceptual explanation. However, the comments might be related with other general aspects of teaching.

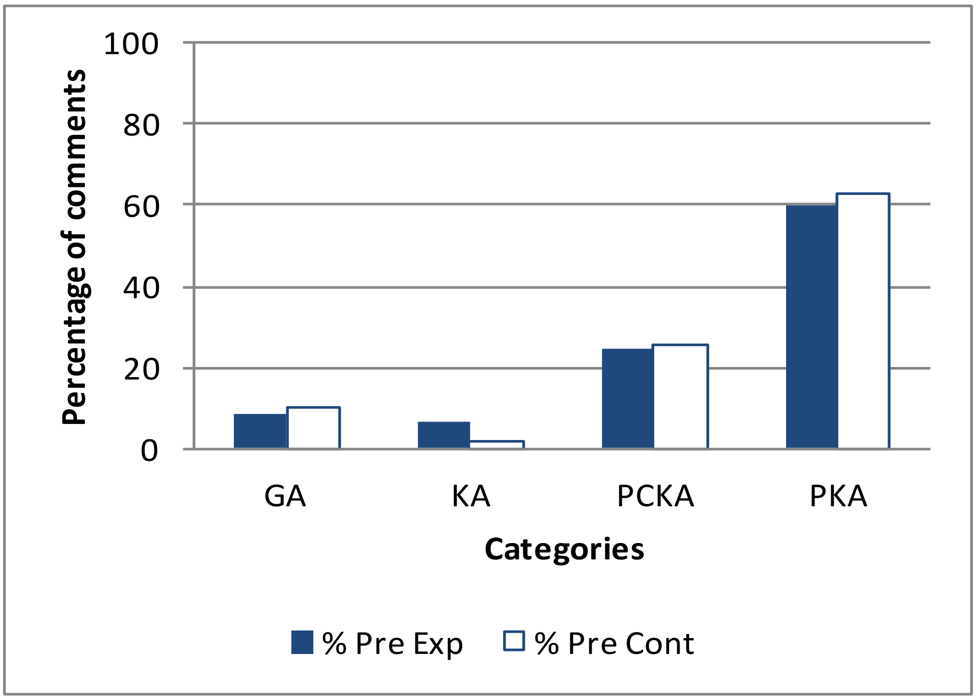
(2) Pedagogical Knowledge Aspects (PKA): Under this category all the comments about relevant issues in teaching through conceptual explanations were grouped together. It shows pedagogical knowledge present in the teachers' mind at the moment of assessing the quality of explanations. However, the comments in this category did not link the pedagogical knowledge with the content the peer teacher was addressing in the explanation.

(3) Pedagogical Content Knowledge Aspects (PCKA): This category refers to all teachers' ideas that teachers had to assess the explanations that showed applied pedagogical knowledge to the content. These ideas meant the student teachers were thinking of the ways to teach more effectively a specific piece of content or how learning that content could be more meaningful, which implies a deeper applied and more flexible knowledge.

(4) Knowledge Aspects (KA): This category includes comments referring to the science knowledge demonstrated in the explanation, about the accuracy of the peer teacher's explanation in terms the usage of scientific terms, algorithms or processes.

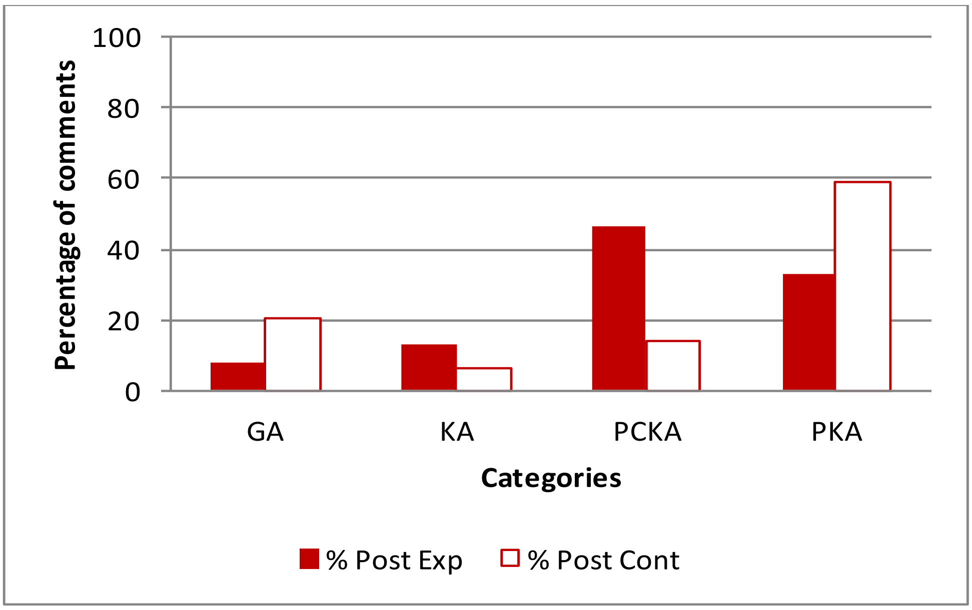
Analysing the justifications that the student teachers from experimental group (Exp.) and control group (Cont.) gave to their evaluation allowed determining if the groups were similar or not in the criteria they used to evaluate the explanation. This measure was important considering the participant group was voluntary. Then, they could have been influenced by other variables of self-selection that might affect their conceptions about the quality of explanations in science. Nevertheless, teachers from control and experimental groups did not present marked differences before PA (Pre measurement), as shown in Figure 7. Indeed, experimental and control groups presented a very similar distribution of their evaluative comments into the categories. Both groups showed most of their conceptions related to PKA, followed by conceptions involving PCKA. In both groups around the 10% of the comments were about GA conceptions and a few were related to KA. This last category was slightly higher in the experimental group.

Figure 7: Distribution of comments in categories Pre PA experimental and control groups



On the contrary, the groups were not similar at the post measurement in their category distribution. The experimental group showed a change in the proportion of evaluative comments in each category, as seen in Figure 8. Most of the comments in the experimental group were related to PCKA, while in the control group the majority of comments continued being related to PKA as it was in the pre measurement. Indeed, in this last group the only modification was a slight decrease of PCKA related comments and an increase of KA related.

Figure 8: Distribution of comments in categories Post PA in experimental and control groups



In the following graphs each group (experimental and control) is illustrated itself, in order to have a better comparison of the change and stability they had respectively. Figure 9 illustrates a distribution change in experimental group comments, while Figure 10 shows more stability in the distribution of control group's comments into the categories.

Figure 9: Distribution of comments in experimental group PRE-POST PA

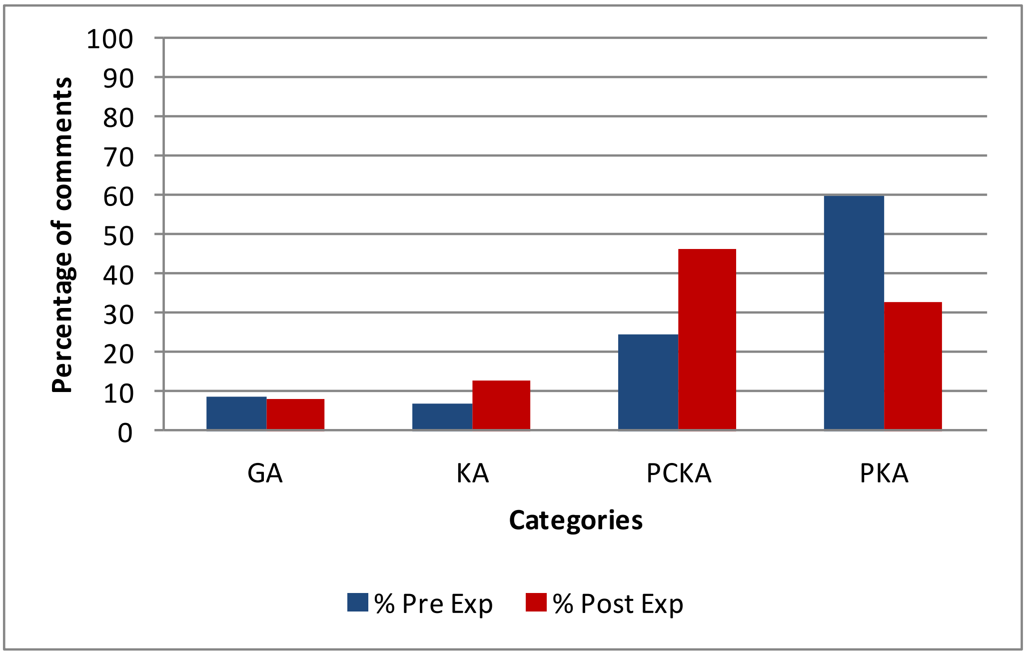
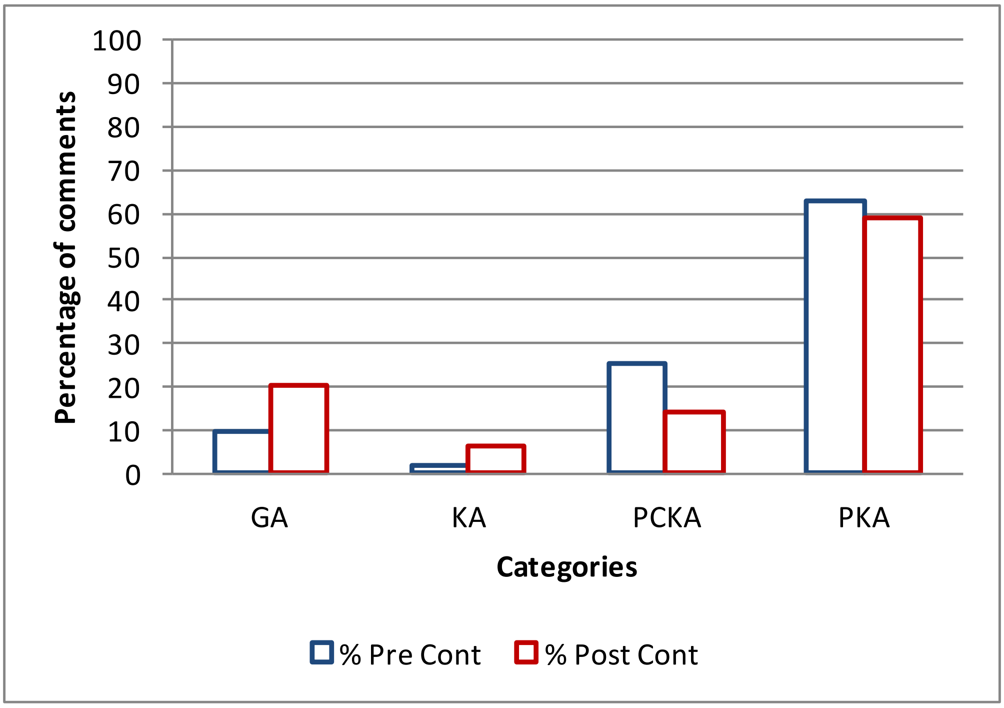


Figure 10: Distribution of comments in control group PRE-POST PA



4.2.3. Pedagogical Content Knowledge development

To determine whether or not the participants' conceptions were different before and after PA in the experimental group, the evaluative comments given as feedback by the teachers during their peers' microteaching episodes were coded in the four categories mentioned in section 4.2.1. Several subcategories emerged from the participants' discourse because of the data richness (exemplified in Appendix 8.9). The categories and subcategories were refined with another researcher. After reaching agreement in their description, inter-rater reliability was calculated double marking the 5% of all data. This was done to estimate how much consensus in the assignation of comments to categories and subcategories existed. The inter-rater reliability was high (81.8%). Table 11 shows the labels of subcategories to identify their membership to the categories, and their frequencies are in brackets.

Table 11: Categories and labels of subcategories emerged in peer feedback sessions

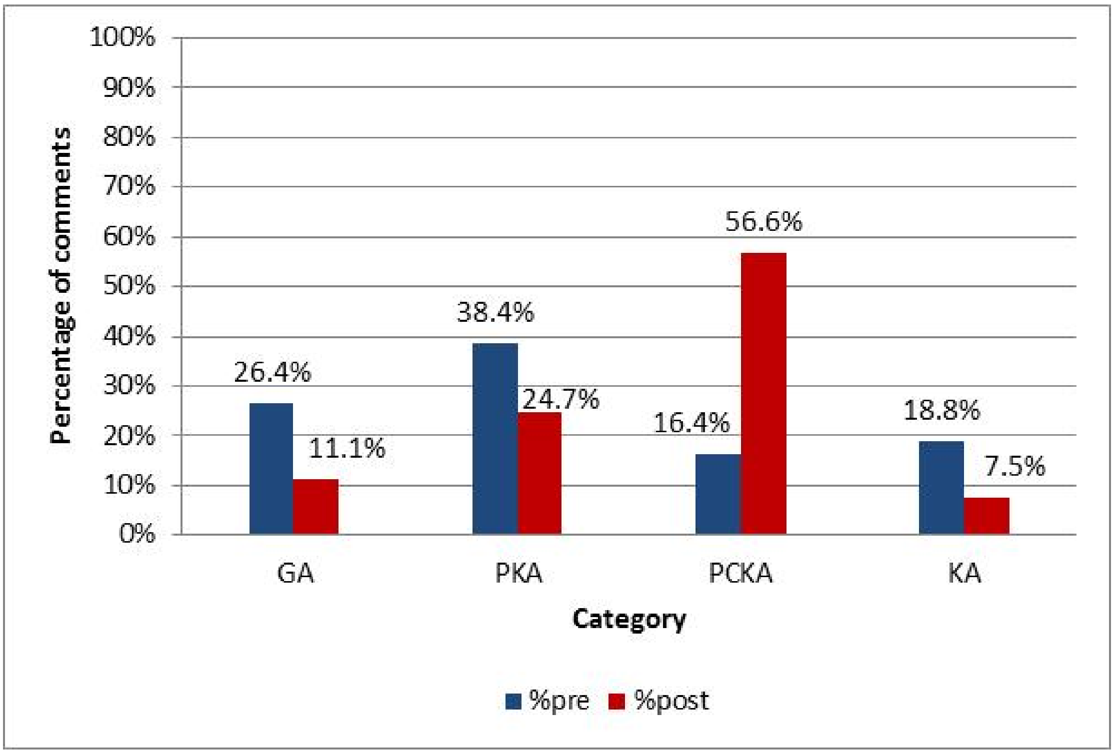
	1.1 Lesson preparation (4)		2.1 Resources availability (3)		3.1 Resource adequacy for content or goal (6)
	1.2 Formal clothes (1)		2.2 Adequacy for pupils' characteristics (18)		3.2 Nature of science (16)
	1.3 Voice, diction and rhythm (25)		2.3 Adequacy for teaching phase (6)		3.3 Analogy accuracy (5)
	1.4 Values in teaching (1)		2.4 Activity goals (8)		3.4 Quality of resource (34)
	1.5 Content v time relation (8)		2.5 Content complexity v pupils' age (12)		3.5 Pupils' ideas integration (15)
	1.6 Excess of content (5)		2.6 Pedagogical language usage (6)		3.6 Mistakes management (6)
	1.7 Clarity of handwriting (9)		2.7 Resource usage (15)		3.7 Content contextualization (13)
1.GA	1.8 Image and (7) speech synchrony		2.8 Pupils' participation (17)	**3.PCKA**	3.8 Connection with other content (11)
	1.9 Importance of contents (3)		2.9 Question type and usage (24)		3.9 Questions' specificity to the content or goal (3)
	1.10 Movement in the classroom (9)	**2.PKA**	2.10 Resource characteristics (18)		3.10 Example adequacy (14)
	1.11 ICTs usage (4)		2.11 Pre-concepts gathering (7)		3.11 Connection between explanation and goal (5)
	1.12 Emphasis on note-taking (2)		2.12 Collective construction of knowledge (7)		3.12 Explanation sufficiency for the content (29)
	1.13 Whiteboard layout (5)		2.13 Questions delivery (8)		3.13 Didactic transposition (44)
	1.14 Lesson title (1)		2.14 Answers' management (6)		3.14 Pupils' ideas transformation (5)
	1.15 Self-confidence (19)		2.15 Schemas title (2)		3.15 Explanation sequence (23)
	1.16 Difficulties overcoming (2)		2.16 Examples usage (10)		3.16 Thinking skills (1)
			2.17 Activity type (1)		4.1 Scientific terms accuracy (61)
			2.18 Gesture usage (8)	**4.KA**	4.2 Up-to-date knowledge (3)
			2.19 Connection with scholar texts (1)		4.3 Processes accuracy (8)
					4.4 Algorithms accuracy (2)

It is interesting to note that important differences were found in the distribution of comments when counting their frequency emerged during the feedback sessions in pre and post microteaching episodes.

As it can be seen in Figure 11, the feedback comments given during the pre-microteaching episodes were mostly related to PKA (38.4%), then GA (26.4%), KA (18.8%) and finally PCKA (16.4%). However, during the post-microteaching episodes the distribution in proportional terms had a very noticeable change. Most of the evaluative comments were related to PCKA (56.6%), followed by PKA (24.7%), GA (11.1%) and KA (7.5%).

This finding is similar to the results found using the PA questionnaires (described in section 4.2.1), implying a reorganization of the emerging categories in teachers' discourse. Both findings suggested that teachers' conceptions changed towards the development of pedagogical content knowledge during PA intervention in the experimental group.

Figure 11: Distribution of feedback comments pre and post into categories in experimental group



4.2.4. Implicit theories about the quality of teacher explanations

The implicit theories the participants held about the quality of teacher explanations were accessed through the construction of an instrument to assess the peers' explanations from the student teachers' views. Two discussion sessions of two hours per group were necessary to explore in their views about the quality of explanations, identify the elements emerging from student teachers' discourse and convert the elements into assessment criteria. The student teachers applied their instrument to assess a selection of explanations in a wide range of quality: video recorded explanations, their peers' explanations and their own explanations, in order to perfect the criteria. All the processes were developed entirely by the student teachers and the researcher played a facilitator role only.

The exercise of constructing a raw instrument implied negotiation the student teachers' understanding of the concept of "quality of explanations" and which characteristics -when explaining and general teaching principles- were the central. The researcher's objective was to make explicit the implicit theories the participants held about the quality of explanations through this activity and product analysis and explore how the meaning was negotiated within the groups.

The three groups (U1, U2 and U3) agreed to assess the quality of the explanations through analysing the process conducted by the peer student, not the explanation as a product only, neither related with students' understanding because of the simulated character of the microteaching episodes.

The instruments generated by the teachers (described and presented in the following pages) were considered in this research as a product of their implicit theories about the quality of explanations. A joint reading of the instrument and its description is suggested, following the numbers in the text that are the number of the criteria in the three instruments.

(a) In the **U1 group**, the implicit theories that embodied the instrument construction (Table 12) were easily associable to the constructivism theory applied to teaching science (Fenshamp et al., 1994). This is a world tendency in science teaching but in Chile it has been difficult to install in teachers' discourse and practice (Cofré, 2010).

This was possible to observe because in the first place the diversity approach appeared (1). The teachers valued in an explanation the explicit inclusion of a topic which was possible to be explained from the diversity approach (Sanmartí, 1994), for instance gender, cultural, ethnical inter or intra individual differences. They also mentioned the diversity approach in their peers' decisions such as the teaching resources selection.

Likewise, how the student teacher handled the content emerged in teachers' explicit discourse, which after questioning in a deeper way let appear the relevance of reaching consensus of the terms (2) used in the explanation between the teacher and the pupils. The student teacher's contextualization of the content (3) appeared as another aspect helping the understanding of concepts. The participant teachers held in the implicit plane that a good context implied putting the content in more concrete, simpler or wider elements, which should be connected with the concept being taught. The participant teachers mentioned its importance because it allows connecting the explanation with what the pupils already know. Besides, the relevance of linking relations between the concepts that are being explained was mentioned (4). The participants implicitly valued working with the link and not only mentioning it to the students, which meant constructing the scientific concept through establishing connections with other concepts. They also indicated those connections need to be correct in terms of science, and useful to support the understanding of the concept. These teachers highlighted linking the explanation with the pupils' everyday life (5), giving value to establishing the differences and similarities between both, and also to how the teacher gathered and used prior knowledge process in the explanation (6). They also mentioned how the teacher used the questions (7) to obtain pupils' answers to integrate or confront them in the explanation (8), valuing the transformation of pupils' ideas in the explanation in a more implicit plane. Finally, in this group of teachers' discourse the use of examples in the explanation (9) emerged, which are useful to illustrate the content when they are pertinent to the content and familiar to pupils' experiences. Also, the emphasis on the pupils' taking notes (10) during the explanation as a way of formalising the knowledge appeared or learned.

Table 12: Student teachers' rubric to assess peers' explanations University 1

Criterion	Low quality (Not achieved)	Medium quality (Half achieved)	High quality (Achieved)
1. Diversity approach: how the teacher explicitly teach topics from a diversity approach.	The teacher does not include in the explanation any topic from the diversity approach.	The teacher includes in the explanation a topic from the diversity approach.	The teacher includes in the explanation a topic from the diversity approach giving examples that globalize it or refer to how the diversity enriches concept understanding
2. Terms usage: How the teacher gives meaning to the concepts.	Most of the terms the teacher uses in the explanation do not have meaning got by consensus.	The teacher gives a definition of the terms without exploring the students' prior knowledge.	The teacher explores in students' prior knowledge about the terms being used, making students participate, correcting their mistakes and potentiating their successes.
3. Contextualization: How the teacher presents a general context to introduce the explanation.	The teacher does not contextualize the explanation.	The teacher asks to the students to contextualize the explanation but does not declare the context.	The teacher contextualizes the explanation in a simple way, interacting with the students and presenting them a concrete context.
4. Link with other concepts: How the teacher links the concept with other scientific concepts.	The teacher does not link the concepts, or the link is conceptually incorrect.	The teacher links two or more concepts, but the link does not support the concept understanding or it is a not clear link.	The teacher establishes a clear and conceptually correct link between two or more concepts, and it supports the concept understanding.
5. Link with everyday life: How the teacher links the concept with elements from the students' everyday life.	The teacher mentions a link with the students' everyday life, but does not explain the link.	The teacher mentions a link between the concept and the students' everyday life but only for a memory function.	The teacher mentions a link between the concept and the students' everyday life mentioning similarities and differences between both without losing the focus.
6. Prior knowledge: How the teacher links the concept with students' knowledge.	The teacher does not gather students' prior knowledge.	The teacher gathers students' prior knowledge but does not use explicitly to explain.	The teacher gathers students' prior knowledge and uses it explicitly to explain, linking it with the concept explained.
7. Questions: How the teacher uses different type of questions and poses them to the class.	The teacher does not ask questions during the explanation or they are always closed.	The teacher ask open and closed questions but poses only to a student or group, or does not wait for the answers.	The teacher asks specific open and closed questions and poses them widely to the class.
8. Answers: How the teacher manages the students' answers.	The teacher does not do anything with the answers or says "good" independently of the answer.	The teacher gathers answers but integrates only the related answers to the question.	The teacher integrates the answers related with the explanation and corrects the errors, clearing the mistakes or allowing students to realise the mistake and self-regulate.
9. Examples: Quality of the examples the teacher gives to explain.	The teacher does not use examples to explain or they are not pertinent to the concept.	The teacher uses ambiguous examples, not familiar to the students or that do not illustrate the concept.	The teacher uses examples pertinent to the content, familiar to the students, accurate and illustrative.
10. Taking notes: Whether or not the teacher encourage students' notes.	The teacher doesn't encourage pupils to take notes during the explanation.	The teacher encourages pupils' notes but does not verify if they do it.	The teacher encourages pupils to take notes and verifies if they do it during the explanation.

(b) The **teachers from U2** created an instrument (Table 13) based on the idea that every explanation constructed for science teaching in classrooms could work as a model of the scientific concept or phenomena being explained, following the ideas of Justi (2006). This group defined eight criteria; three of them described the explanation usage in the wider lesson context and the other five defined its quality.

The first three criteria identified the moment when the explanation appeared (1), at the beginning, middle, end or in different parts of the lesson. These student teachers also indicated in their rubric the observable function the teacher gave to the explanation (2), such as motivational, demonstrative, explanatory or evaluative, and the percentage of the lesson time (3) the teacher used to explain.

Within the quality criteria they mentioned -as the U1 group- the links with pupils' prior knowledge (4) had priority at the moment of explaining. The implicit idea of this criterion was connecting pupils' ideas with the proposed model of the scientific concept is positive only when the teacher explicitly uses the prior knowledge in the explanation. Likewise, for these student teachers the explicit emphasis on the representative character of the explanation as a model was a determining aspect in the quality of the explanation (5). The implicit theory here was there are many ways of representing knowledge and the explanation is just one of them. Thus, teacher communicates the nature of science through making visible the distance from representation to reality.

The participation of the pupils in the explanation (6) was another important criterion for this group. It elicited student teachers' views about the constructive process of explanations in science, which was flexible to integrate pupils' questions, ideas, etc. The accuracy of the teacher explanation (7) also appeared in this group of student teachers' discourse. However, going deeper in the implicit plane this group thought that a teacher who explains correctly and answers all the pupils' questions is better than the one who is explaining correctly but leaving unanswered questions. Besides, this group of teachers highlighted the importance of the clarity of the explanation (8), which was connected implicitly for them with the conceptual clarity the teacher had about the scientific concept. In their ideas, if the teacher has clarity about the content knowledge is not making mistakes when explaining and this enriches the use given to the explanatory model. Otherwise, if the teacher does not have a clear content knowledge, giving a clear explanation through the model will be not possible.

Table 13: Student teachers' rubric to assess peers' explanations University 2

Criterion	Indicator or level
1. Lesson moment where the teacher explains.	Beginning: The teacher explains at the beginning of the lesson.
	Middle: The teacher explains at the middle of the lesson.
	End: The teacher explains at the end of the lesson.
	The entire lesson: The teacher explains in different parts of the lesson: at the beginning, middle and or end of the lesson.
2. Observable function the teacher assigns to the explanation trough a model.	Motivational: The teacher promotes the students' motivation.
	Demonstrative: The teacher explains nature elements through examples.
	Explanatory: The teacher explains phenomena or processes that occur in nature.
	Evaluative: The teacher evaluates students' knowledge to challenge their prior theoretical knowledge.
	Another: *Detail here please*
3. Percentage of time used by the teacher to explain.	◯ 0 – 10% ◯ 11 – 20% ◯ 21 – 30% ◯ 31 – 40% ◯ 41 – 50%
	◯ 51 – 60% ◯ 61 – 70% ◯ 71 – 80% ◯ 81 – 90% ◯ 91 – 100%
4. The teacher gathers prior knowledge from the students and integrates it with the explanation.	Not achieved: The teacher neither gathers nor identifies the students' prior knowledge about the content or the model presented.
	Half achieved: The teacher gathers and or identifies the students' prior knowledge about the content or the model presented, without linking them with the model.
	Achieved: The teacher gathers and or identifies the students' prior knowledge about the content and links explicitly the prior ideas with the explanation or model.
5. The teacher refers the model used to explain is a representation of the reality and there are others possible.	Not achieved: The teacher does not refer implicitly or explicitly the model used to explain is a representation of the reality and, but assumes the model is the reality.
	Half achieved: The teacher refers implicitly or explicitly the model used to explain is a representation, without mentioning implicitly or explicitly the existence of other models to explain, or that it is a provisional model.
	Achieved: The teacher refers implicitly or explicitly the model used to explain is a representation of reality and there are other models to represent the content.
6. The teacher makes students interact with the explanation.	Not achieved: The teacher does not make students interact with the explanation.
	Half achieved: The teacher achieves partial interaction between the students and the model, because there are doubts about the explanation and its uses.
	Achieved: The teacher achieves interaction between the students and the model through the students' participation in the explanation of the model or questions.
7. The teacher explains correctly the model proposed.	Not achieved: The teacher does not explains correctly, causing conceptual mistakes in the students.
	Half achieved: The teacher explains correctly, but making mistakes when answering students' questions, or the teacher does not answer all the question
	Achieved: The teacher explains correctly and answers all the questions raised from the students.
8. The teacher demonstrates explanation clarity when explaining the model.	Not achieved: The teacher does not have a conceptual clarity, which causes making mistakes when using the model.
	Half achieved: The teacher has a medium clarity about the concept being explained at the moment of using the model.
	Achieved: The teacher has plenty clarity about the content being taught, which enhances the usage of the model.

(c) In the **U3 group**, the analysis of the construction of their instrument (Table 14) indicated it was possible to assume they had a simpler view about explanations in science. A few elements were similar to the other two groups of teachers, but teachers from U3 presented less sophisticated ideas which were more difficult to transform into criteria.

For these teachers the use of examples in the explanation (1) was the most important element to define its quality. After questioning about the characteristics the examples should have and applying the criteria, it was observed that the good examples for them were: familiar or close to pupils' experience, as concrete as possible and related with the scientific concept being explained. The connection with pupils' prior knowledge (2) emerged in the second frequency, which was also found in the other two groups. For U3 teachers the clue was gathering what students already know trough questions, and linking this knowledge with the concept being explained. However, the questions (3) should also have a quality aspect; being posed to the entire lesson without giving priority to one student or a group of them for particular reasons.

A different aspect of explanations that appeared in this group and not in the others was the sequence and concision of the explanation (4). They mentioned the explanation should have neither unnecessary nor missed elements, and it must have a connective thread. Exploring deeply in this idea, the implicit theory appeared making the connection between both aspects; if there are missed elements the thread would be broken, and only if each part of the explanation conducts to another, a good sequence would be established. Otherwise, isolated elements or not connected with others, would be unnecessary parts for the explanation.

In terms of the accuracy of the explanation (5), the teachers asserted that the teacher must handle content knowledge about what is explained. The way in which they implicitly referred to this was is in the precision of the explanation or when the teacher is not redounding in some aspects, because redundancy meant for them the teacher is staying only in his 'safe place'. Another related element was what the teacher does with students' answers (6). They mentioned clearing the conceptual mistakes and integrating it in the explanation as relevant actions. Nevertheless, in their discourse the teacher needs to have good content knowledge to be able to correct students' misconceptions. Then, both criteria were clearly connected.

Finally, this group mentioned the collaborative work (7) as an important criterion in the quality of conceptual explanations. By collaborative they referred to constructing the explanation between the teacher and the pupils and also between the pupils. It could be achieved through activities that allow collaborating, which reflects a more flexible view about the nature of the science knowledge and its construction.

Table 14: Student teachers' rubric to assess peers' explanations University 3

Criterion	Low quality (Not achieved)	Medium quality (Half achieved)	High quality (Achieved)
1. Examples usage: Quality of the examples the teacher gives when explaining.	The teacher does not use examples when explain or the examples used are not related with the concept being explained	The teacher uses concrete examples that are related with the concept, but they are not close to student's experience or knowledge.	The teacher uses concrete examples that are related with the concept and are close to student's experience
2. Prior knowledge: How the teacher relates the concept being explained with the students' prior knowledge.	The teacher does not gather students' prior knowledge or ideas.	The teacher gathers students' prior knowledge or ideas but does not use them explicitly to explain.	The teacher gathers students' prior knowledge or ideas and uses it explicitly to explain, linking them with the concept.
3. Questions: How the teacher different type of questions and poses them to the class.	The teacher does not ask any question during the explanation.	The teacher opens a moment to ask questions (open and closed), but they are directed only to a student or a group.	The teacher opens a moment to ask questions, directing them widely to the students.
4. Sequence and concision.	There is not a conductive tie in the explanation, or it is interrupted because more than one part of the explanation is missed or unnecessary.	Each part of the explanation conducts to the next one (conductive tie), but there is one part of the explanation missed or unnecessary.	Each part of the explanation conducts to the next one (conductive tie), and there is any part of the explanation missed or unnecessary.
5. Accuracy/Concept handling.	The teacher does not handle the concepts being explained, there is redundancy, mistakes or he induces conceptual mistakes in the students.	The teacher handles the basic concepts, but when explaining is not accurate (there are inaccuracies).	The teacher demonstrates handling the concepts because the explanation is accurate and there are not mistakes.
6. Answers management: What the teachers does with the students' answers.	The teacher does not do anything with students' answers or says "good" independently of the quality of the answer.	The teacher integrates only the answers that seem correct for him, or does not correct the inaccuracies in the student's answers (they keep the mistake).	The teacher integrates the answers related with the explanation and corrects the errors, clearing the conceptual mistakes.
7. Collective work with concepts.	The teacher does not do any type of collective work with the concepts.	The teacher works collectively a concept.	The teacher works collectively a concept, giving it a collective meaning.

To summarise, the construction of an instrument as a device to explore the participant teachers' theories (implicit and explicit) was considered highly valuable for the purposes of this research. Through this process of product analysis it was possible to observe the group of student teachers' theories varied according to the university they belonged to, then, probably the variations were explained in part by the different science knowledge they had, and perhaps by the programme transmitted views.

Although all the groups defined themselves as believing in a constructivist way of teaching science, at the moment of deciding why an explanation was better or not, the quantity of elements related to constructivist approach were very different. In this sense, U1 and U2 presented more elements than U3, and in this university the implicit theories about quality of explanations were including broader elements, not only useful to analyse and assess explanations of scientific concepts, i.e. answer management, collective work.

Nonetheless, there were two common points between the three groups of student teachers; the use of examples and the interaction between the teacher and the pupils during the explanation. Those aspects were the most mentioned in participant teachers' discourse, and they were also the highest developed or easier to change as indicated in section 4.1.2. However, in any groups' implicit theories appeared elements such as analogies, metaphors or simulations, or using mistakes as a learning opportunity. Those were also the lowest developed and the most difficult to change. Then, a possible connection could be established by the researcher, showing implicit theories teachers held about the quality of explanations as guiding teachers' practice to explain scientific concepts.

4.2.5. Quality of student teachers' explanations

The quality of the explanations of teachers who participated in PA was measured in the initial microteaching episodes (before PA intervention) and the final microteaching episodes (at the end of PA intervention). These episodes were assessed with the rubric (in Appendix 8.8), which contained ten criteria and a final score from 0 to 20 points as mentioned before. In every statistical analysis Levene's test was run to allow further calculations.

In the description of this section the participants were divided into the three original experimental groups (U1, U2, and U3) as they belonged to three different universities. As indicated in Chapter 3, the universities differed in their students' knowledge of science (high, medium, low), measured by the number of science courses they required the pre-service teachers to graduate. U1 gave their students fourteen courses, U2 nine, and U3 four courses.

Pre-test

In the initial microteaching episodes the student teachers presented a varied quality of their explanations. Although there were some student teachers with a high score, the general pattern showed a medium performance as marked against the rubric. This suggests most of the items were half achieved. According to the totals, the minimum score in quality of explanation was 5 and the maximum 17 with standard deviation 3.38 (see Table 15).

Table 15: Descriptive Statistics experimental group PRE-PA

Group	Mean	N	Std. Deviation	Minimum	Maximum
U1	9.50	6	4.03	5	14
U2	11.25	8	3.28	6	17
U3	9.17	6	2.92	6	14
Whole group	10.10	20	3.38	5	17

Considering that the universities were selected according to science knowledge variation criterion, it was expected U1 would have shown a better performance than U2 and U3 because their student teachers had a higher knowledge about science. However, student teachers from U1 shown 9.5 points in average, which is very similar to the score obtained by teachers from the university with the lowest science knowledge (U3, 9.17 points). Surprisingly for the researcher, U2 presented a better score average in the initial quality of explanation (11.25). This institution had an integrated curriculum with science and education courses in similar proportions. It was taken in this research as medium in the science knowledge variable.

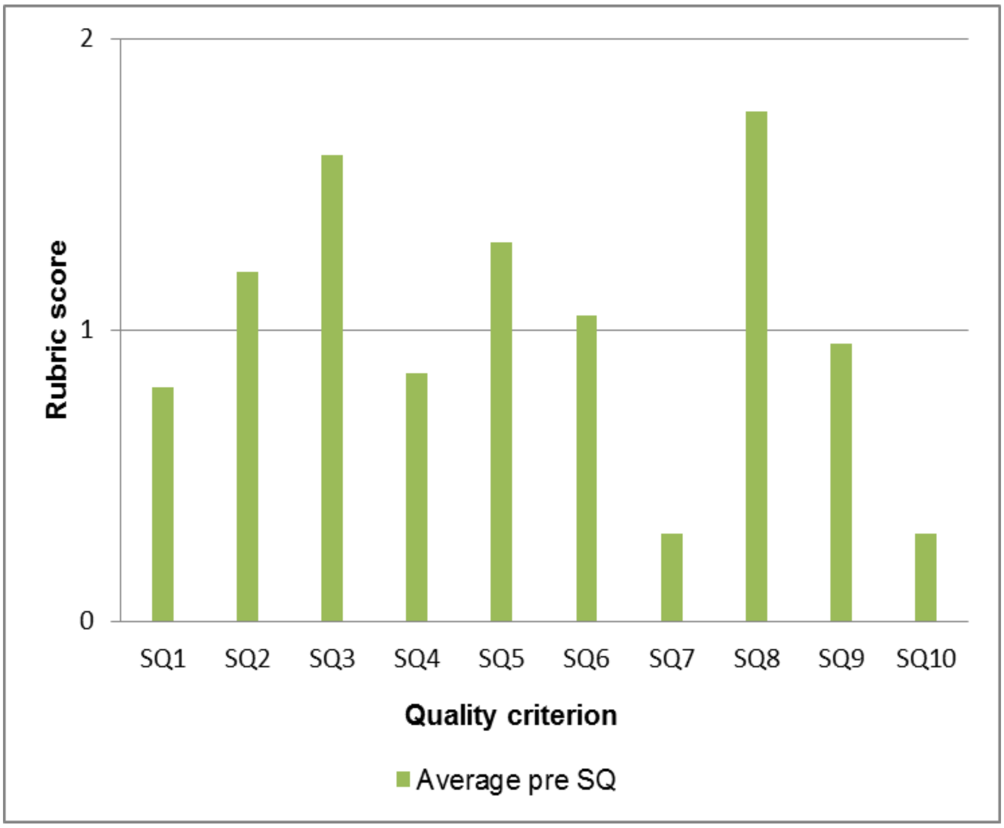
In order to explore if this difference was statistically significant, a one-way ANOVA was run, but the differences comparing the three groups were not statistically significant (df=19, F=0.764, p=.384). A Student t-test for independent samples was also run, grouping the values of teachers from U1 and U3, but again the differences were not statistically significant (df=18, t= -1.259, p=.224). Then, it was assumed the groups had similar quality of their explanations in the pre measurement.

In terms of heterogeneity within the groups, it was possible to assert from the standard deviations that all the groups were quite heterogeneous within themselves but somewhat similar to each other. Each group had student teachers with a low and medium performance. In U2 only, there was one student teacher with an initial score that could be considered of a high level (17 points of a maximum of 20). These results suggest that at the end of initial teacher training, the quality of this group of student teachers' explanations according to the rubric was not directly associated with the number of science courses they have taken during the previous years. Likewise, the results of U2 could imply that having a balanced combination of integrated science courses and education courses may help pre-service teachers to explain science better. This is a researcher's interpretation.

The rubric allowed a detailed view about the specific areas where the teachers had weaknesses or strengths. The items were: (SQ1) clarity; (SQ 2) coherence and cohesion; (SQ 3) sequence; (SQ 4) accuracy; (SQ 5) sufficiency; (SQ 6) connection with pupils' experience; (SQ 7) usage of analogies, metaphors, simulations or models; (SQ 8) usage of visual representations; (SQ 9) usage of non-verbal language and (SQ 10) usage of pupils' mistakes or common mistakes as a learning opportunity.

Thus, it was interesting to report the results of the participant teachers in the instrument items. Although a numeric comparison is not the best when the data are based on elements absence or presence, each rubric criterion could have an individual score between 0 and 2 depending on its quality. For comparative purposes, the whole groups' scores are shown in Figure 12 in the next page.

Figure 12: Rubric average score by quality criterion PRE-PA



During the initial microteaching episodes the group of participants presented a heterogeneous pattern, with high and low scores in the quality criteria (SQ in Figure 12). On one hand, the rubric items that were the most developed in this group of teachers were the sequence of the explanation (SQ3) and the use of examples, graphs, images or demonstrations (SQ8). On the other hand, the items that the group showed weaker were using metaphors, analogies, simulations or models to explain (SQ7) and using the error or common mistakes as a learning opportunity (SQ10). The other six criteria were considered to be in a medium level rated against the rubric.

It is important to consider that having an average score lower than one (SQ1: clarity, SQ4: accuracy, SQ9: use of gestures and voice) could suggest there were more teachers who did not present the criteria in their explanation, than teachers who presented it in a high with quality level. Likewise, the criteria with an average score higher than one (SQ2: coherence and cohesion, SQ5: sufficiency, SQ6; connection with students' experience), may indicate there were more teachers who presented the quality criterion in their explanation but not with the expected quality than teachers who did not present it.

Post-test

It was expected that after PA there would be a general improvement in participant teachers' performance, measured in their post-test microteaching episode explanations. In general terms in the final microteaching episodes the student teachers presented a wide spectrum in the quality of their explanations. The minimum score in the whole group was 6, the maximum was 19, and the standard deviation was 3.38 as shown in Table 16.

Table 16: Descriptive Statistics experimental group POST-PA

Group	Mean	N	Std. Deviation	Minimum	Maximum
U1	12.83	6	2.92	9	16
U2	16.50	8	2.33	12	19
U3	13.33	6	3.98	6	17
Whole group	14.45	20	3.38	6	19

Student teachers from U1 achieved 12.83 points in average, which is very similar to the score obtained by U3 teachers (13.33). Student teachers from U2 presented the highest overall average score (16.5). This trend was the same at the beginning of the PA intervention, and for this reason it was decided to run the same analysis; one-way ANOVA to compare the three groups and Student's t-test for independent samples with U1 and U3 grouped together. The means difference between the groups was not statistically significant when comparing the three university groups (df=19, F=3.013, p=.76). However, it was statistically significant when comparing the aggregation of U1 and U3 against U2 (df=18 t=-2.504, p<.05) in a Student t-test for independent samples, although the Effect Size was considered small (ES=0.2) (Table 17 and 18).

Table 17: Student t-test for independent samples aggregated groups (U1+U3) vs. U2

t-test for Equality of Means						
t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
					Upper	Lower
-2.504	18	.022	-3.417	1.365	-6.284	-.549

Table 18: Descriptive Statistics aggregated groups (U1+U3) vs. U2

	Group	N	Mean	Std. Deviation	Std. Error Mean
SQtotal	U1+U3	12	13.08	3.343	.965
	U2	8	16.50	2.330	.824

Comparing the pre and post test scores, the student teachers had a higher overall post-test score compared with themselves in the pre-test (mean pre-test=10.1; mean post-test=14.65; SD =3.38). This difference reached 4.55 points in average. The Effect Size was high (d=1.34) according to Cohen's (1988). When a one-way ANOVA was run (see Table 19), there was a significant difference between the pre and post conditions (F=16.54, df=39, p<.001). Then, it is possible to state the improvement in the quality of the student teachers' explanations after PA was statistically significant compared to the initial quality of their explanations.

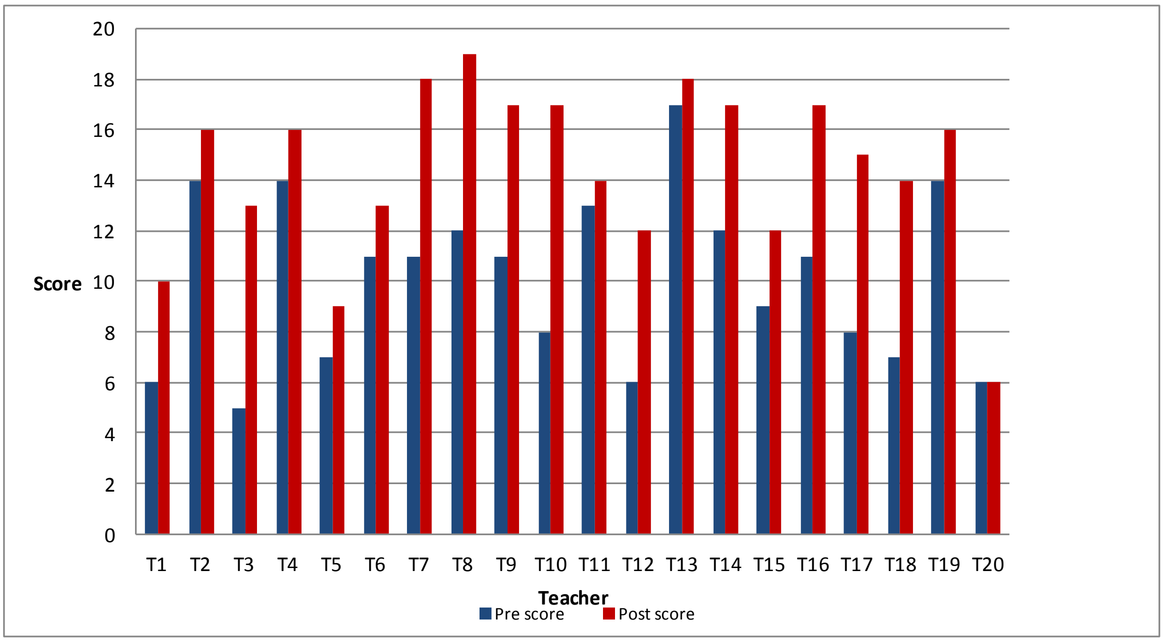
Table 19: ANOVA one way PRE-POST (whole experimental group)

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	189.225	1	189.225	16.540	.000
Within Groups	434.750	38	11.441		
Total	623.975	39			

Looking at case to case (Figure 13), it is noticeable that only two teachers (T5, and T20) showed a small or no advance and stayed below the median line. In general, teachers who started at a lower level improved most at the end of PA.

This could mean PA is especially useful to work with teachers who presented initial difficulties. The highest advances were from teachers from the three universities (T3, T7, T8, T17, and T18), which might imply PA worked independently from their science knowledge.

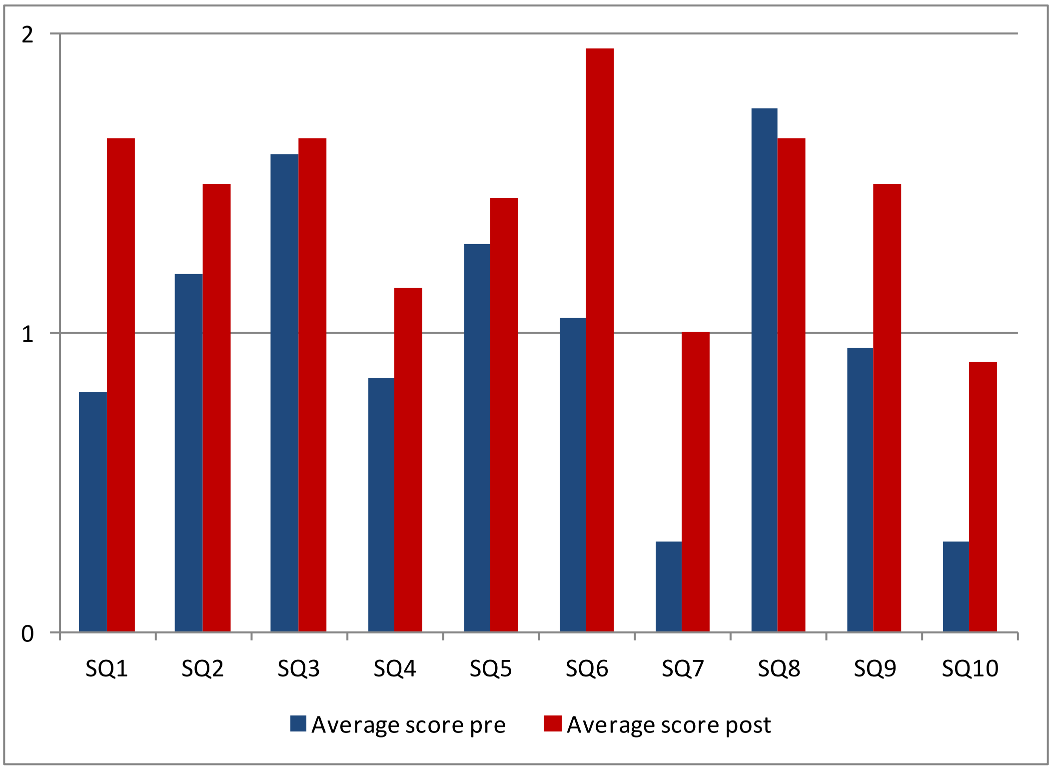
Figure 13: Quality of explanation per teacher PRE and POST- PA



Regarding the general pattern, as Figure 14 shows, most of the rubric items presented a higher score in the post measurement than in the pre measurement. Comparing the pre and post patterns, it is observable that the two lowest criteria in the post score were the same as in the pre score (SQ7, SQ10). This implies although it was possible to have an improvement, this was not as large as necessary to obtain high quality in those aspects.

At pre-test, the student teachers virtually did not use metaphors, analogies, models, simulations (SQ7) or the illustration of pupils' mistakes (SQ10) as learning opportunity. At the final phase of PA although some student teachers used them, this was not with the expected quality. Probably both elements require the development of a mature or deep science content knowledge understanding, which the pre-service teachers do not have yet. Also, it is important to note that this kind of knowledge development was not the focus of PA intervention. These criteria probably need more time and focused interventions to be modified. Besides, in Figure 14 it is noticeable not all criteria were equally easy to change during PA intervention.

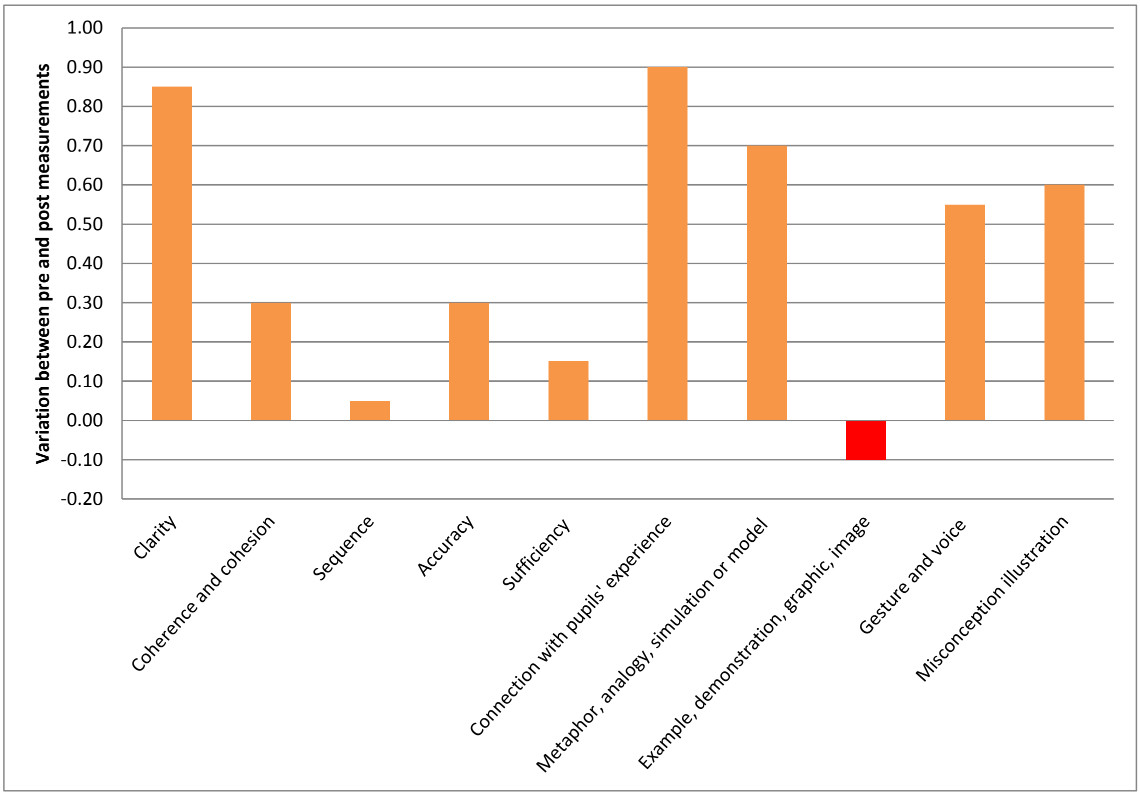
Figure 14: Average score in rubric PRE and POST-PA by quality criterion



To have a clearer view about the criteria modifiability, Figure 15 shows the variation between pre and post scores by quality criterion. The easiest criteria to improve using PA in initial science teacher education were clarity (SQ1) and connection with pupils' experiences (SQ6). This improvement was very marked. It meant that at after PA most of the teachers developed clear explanations and they were able to relate them to their students' experiences, daily life or ideas, which is a main principle in constructivism applied to science teaching (Fenshamp et al., 1994).

Otherwise, there were teaching elements difficult to improve using this methodology. Indeed, there was a criterion that did not improve. The use of examples, graphs, images and demonstrations went down slightly in the post-test. Although it was not expected to see a large improvement because this criterion was one of the highest scored in the pre measurement, the teachers reported that after discussing about the characteristics of a good example, some of them felt less confident to do it in the final microteaching episode. This might explain its decrease.

Figure 15: Variation PRE and POST-PA by quality criterion



4.2.6. Elements promoting changes in teachers' conceptions and practices

As it was mentioned previously, the experimental group that received PA showed changes in their conceptions about quality in explanations and also in their explaining skill rated against the rubric. Thus, it was interesting to explore in the elements that promoted the changes, viewed from the participants' perspective and combined with the researchers' analysis.

Three focus groups were carried out with almost all the participants (18 of 20), one meeting per each university group. In the U1 five student teachers from the six attended. In the U2 group seven from the eight student teachers attended. In the U3 groups the six student teachers participated in the meeting. The focus group question guide (Appendix 8.7) intended to investigate the perceived changes, the elements associated and their consequences.

The participants' ideas were organized in the six categories corresponding to the selective coding proposed by the Grounded Theory by Glaser and Strauss (1967) (central phenomenon, causal conditions, contextual conditions, intervening conditions, strategies or actions and consequences), but the analysis followed the three coding processes; open, axial and selective. The map of codes used to analyse the focus group is presented in Appendix 8.10.

The subcategories were matched within the six categories to construct a model of the phenomenon (Figure 16 in page 142) combining the teachers' ideas and the researcher's analysis. Then, the model goes beyond the data. It is described and exemplified in the next pages, identifying after each example in brackets the Focus group number (F1, F2, etc.), followed by the teacher ID (T1, T2, etc.), and the paragraph number from which the quote was extracted.

The process of changing teachers' theories about the quality of science teacher explanations was set as a **central phenomenon**. These could be implicit theories or explicit conceptions student teachers held. The change was described as a process by seven student teachers:

> *And everybody points to the conceptual change, and the conceptual change is not black or white, our ideas were changing gradually. (F1, T5:23)*

Before, I thought that a good quality explanation was for example, when you know everything by memory, you can stand up in front of the class, and recall the topic transmitting it. But I never thought about giving emphasis to the use of examples, to a correct use of the concepts and prioritize them well … to make them easier to understand and to be more consistent about that, because the examples must be consistent with the concept, about what you are teaching, and then you are creating the model. (F3, T20:8)

To understand the changing process, the researcher organised the **causal conditions** of the phenomenon as facilitators, because teachers mentioned them answering to the question why the change happened? These are the five causal conditions:

Three participants mentioned reflection on their own practice as a cause of their change, which allowed them confronting the theory they had with the practice they were able to perform during the microteaching episodes. Six student teachers remarked the importance of this point because they had little or any teaching experience before participating in the PA intervention, then this was one of their first opportunities to put in practice what they thought about teaching.

I would consider that here teacher reflection process is much more valuable that how he made the lesson. (F1, T6:39)

And because each of us says we follow the constructivism but in the practice, but when you stand in front of the class it is different. (F1, T2:30)

For example at the beginning I think we did it in September, so we had little experience practising as teachers. And we had not seen each other as teachers yet. (F2, T9:32)

In the reflection on their own practice during the PA intervention, three student teachers valued the opportunity of making visible their strengths and weaknesses, as shown here:

Because sometimes we make mistakes, and we see them, but we do not make them visible. (F1, T6:1) And not necessarily because I did not realize they were wrong. But because I am afraid to admit I was wrong because of the social impact it has. That is extremely strong. (F1, T6:1)

I think it is very important to know your own weaknesses. Because knowing what we're missing, we can work to improve it while maintaining the strengths. (F2, T7:38)

Now I am taking certain points and they could be taken step by step and realize that ok, she is failing in this, we have to improve it. In this she is fine, so we must strength it. I think that's the main point. (F3, T20:13)

In this reflective process, the researcher noticed that when teachers assessed their peer's practice they were moved by psychological mechanisms of projection and reflection. As in a mirror, the peers' practice worked as a reflex of what pre-service teachers could do in a similar teaching situation. Likewise, the teachers who were in the assessor role were identified with the practice and imaginarily projecting their own possible decision making process. These mechanisms were seen in comments from seven of the twenty participant teachers. Some examples are shown as follows:

Reflecting: The assessed teacher reflected real or potential teaching situations or tensions he has faced or would face, which are recognised by the assessors.

> *I meant when the teacher spoke of the matter. That also happened to me! (F1:T5, 31)*
>
> *But creating the reflection space, in this case I saw what I did not have: interacting with the students. (F1, T5:56)*

Projecting: The assessor teacher projected his own field of pedagogical decisions or teaching behaviours on the peers' imaginary decisions and the current practice.

> *And this allowed us to think "but if I would have been explaining it, how I would do it? How would I take it?"(F1, T4:44)*
>
> *It's different now because we put in the other's place ... It helps us put ourselves in the place of the other, what he wanted to accomplish or what was he expecting with the thing he did, and without evaluating ourselves it would not have been possible. (F2, T7:51)*

Also it is possible to observe the convergence of both mechanisms in this teacher' quote:

> *I think we can easily drop to criticize a lesson, saying 'you know, this is good and this is wrong'. But when you start comparing yourself with, seeing your own lessons, you could say 'you know that I made the same mistakes but I did not realize'. And from this work you can say 'maybe I should have used another concept, or I should have done something else.' (F2, T13:35)*

The last element categorised as a causal condition was the modification of the student teachers' focus of analysis, which was seen in comments from twelve student teachers. It means, what student teachers focused on to evaluate the quality of an explanation was modified, and also their analytical view was more critical, including positive and negative points of the peers' performance as shown here:

I remember the first lesson, I was saying "no, this is bad ..." and our analysis was all negative. But after that, with the sessions advancing, it was not everything negative. We did not ignore it but we acquired the habit of saying "you did this wrong but also there were other good things". Then, this is good at the moment of assessing. (F1, T4:43)

I put the same mark but I feel I changed my opinion. Now I'm more critical. I was more critical in what I observed. And also in the last part I detailed more the possible explanation I would do, I took more into account how I was modelling it, to make children realize that it was a representation. And I was not like that at the beginning. At the first moment it was more superficial the critique I gave, now I think it is more sufficient, deeper. (F2, T10:11-12)

I initially looked at what happened to the children, and now I focused more on what happened to the teacher. Then I was completely out of focus in the first activity. And about the teacher ... now I focus on what she decided to do, or what she did wrong, etc. (F2, T7:19)

I agree on the fact that now we are more concise and precise in what we are assessing, we are not focused only in macro aspects, but we are more focused now in the connecting ties of the lesson, or things that we were not focused on before. Then, now we can attack direct points, not general. (F3, T15:16)

Now I realize, for example, ok, she handles the concepts... I do not know, I used to write things like 'the lesson is boring', not that there were not teaching resources ... But now I am focused in other things, like the extent to which she domains the content, or her use of teaching resources. (F3, T17:12)

It is important to mention that from student teachers' perspective having an instance to perform teaching practice was crucial. It could be simulated in a protected context like in PA intervention or in a recorded lesson with real students. The student teachers' ideas implied that counting with real teaching material was necessary for the other four causal conditions happen: visibility of weaknesses and strengths, reflection-projection, confronting theory-practice, and changing analysis focus. This is why it is in the centre of Figure 16 (in p. 133).

Much! And also it was important because last year we did not have our period of practice yet, then we have not teach lessons. We only had observed or had done some small interventions. But now here we performed lessons, then we have experience now. Just a few but we have, and we can take this teaching experience to improve it, or to fix certain things. (F2, T7:34)

Nevertheless, important **contextual conditions** allowed the reflection on practice elements work together, and in this research they were considered also facilitators.

Five participants valued the critique as a constructive tool when given respectfully:

> *What I valued the most was the critique in a respectful framework. (F1, T6:1)*
>
> *And now, it takes shape, because here I tied together the constructive aspect with the critique and how you can make an improvement. (F3, T20:19)*

Likewise, five student teachers noticed the interchange of roles between being assessed and assessor was a very important characteristic of PA, which could be a complex process at the beginning but finally beneficial. This is seen in the following comments:

> *At first you think it is the same to be evaluator than being evaluated, and it is complex, it is more difficult at the beginning. (F1, T3:3)*
>
> *We observed ourselves in this work, and we also could observe other teachers, as a third party, and we could identify it from what we designed... these criteria... they are possible to observe in the practice! This also made us grow professionally. (F2, T8:30)*
>
> *I think so, it's a good way to learn, evaluating yourself and evaluating others. (F3, T20:68)*

This role interchange was facilitated according to one of the participants because the teachers shared the same level of prior experience teaching, knowledge about science and teaching and they were in a similar age range. It is useful to remember that assessment groups were conformed within the same university, to enhance this peer recognition between participants.

> *Among peers it is closer [the assessment], because there is not a situation of hierarchy, because it is more horizontal, and closer in the conversation too. Also here we all had the same level, theoretical and practical.... And our informal language maybe allowed that. (F1, T6:1)*

Other group of factors that teachers identified as important in their process of change were categorised as **intervening conditions**, which had a direct impact on the reflection on practice process. These appear in the Figure 16 as arrows.

The first intervening condition was the systematic evaluation of student teachers' practice. Three participants asserted that doing microteaching episodes and being systematically evaluated helped them, and it could be also beneficial if repeated in their placement centres.

> *Furthermore, the fact that we have made the microteaching episodes helped us to realize in which aspects we were failing, and you could take some points where are the weaknesses. (F3, T20:22)*
>
> *This would benefit also in our own placement centres, because from my point of view they assess one session only, and sometimes it does not reflect the entire process of teaching. (F1, T5:58)*

Secondly, one of the student teachers valued the empathy of the PA facilitator:

> *It is noticeable that she not only does it because she has to complete the task or because she must cover certain things, but she does so in a human way and this is also much appreciated. (F1, T4:44)*

Working with the rubric that student teachers created to analyse their peers' practice was meaningful for four of them. They recognised it allowed them to know new criteria and it was valuable when being evaluated, because they knew the assessment criteria in advance:

> *So I think having the assessment parameters it helps a lot, because I know what my classmates are looking for: this, this and this. Then that is useful to outline the lesson. (F3, T15:29).*

Taking into account the contextual conditions, plus the causal and the intervenient ones it is possible to picture how the elements promoting changes were related. The causal factors were given within this specific context only, and the intervenient conditions were mediators that facilitated causal factors acting together to allow teachers' reflection on practice. When this setting was established in the frame of PA, student teachers were able to generate two **strategies** that gear the changing theories process: the negotiation of meaning and self-regulation to improve the practice.

Constructing the criteria collectively from their own views about teaching was important for three of the student teachers, because the criteria were agreed through a consensus. This reaching consensus process was the key to negotiate the meaning about the quality of science teacher explanations. Its importance can be seen in the following comments:

> *But the main thing here is that the criteria were made from group consensus. And I think the consensus on any assessment is important, and in this aspect we were fine. (F1, T3:5)*

> *And I liked that we had the opportunity to speak about every point of the rubric, because we discussed them, clarified them, modified the criteria, and we were not aware about that before. (F2, T10:26)*

The self-regulation strategy was understood as the decision making process in which the participants took decisions about their own teaching practice, based on the self-analysis and awareness. This was observed in the comments from three student teachers:

> *Also what you do here is regulating yourself… I did not know that I was wrong here, then, I will get better. (F1, T1:33)*

> *We must improve things like that, and we need to analyse ourselves and be aware of that. (F3, T20:8)*

Their self-regulation allowed student teachers making concrete changes in their practices in order to improve it. These changes were described by seven student teachers. From the researcher's point of view, it was the immediate consequence of the changing theories process, also fed by the negotiation of meaning about quality of science teacher explanations.

> *Then you realize, you appropriate this mistake, and you are able to modify it and not do it again. I am more cautious now. (F2, T13:36)*

> *For example, I was told that … they did not correct me but they let me know I could do it better than I did it the first time. I tried to improve it now, by involving more students in the explanation, moving the class more, and I think I succeeded. And that I feel that comparing what I did this time with what I did now, I think it is much better. I made some inaccuracies but I generated a better explanation, more concise. Even without graphic support it was more composed, more conformed. (F3, T15:28)*

> *I think we improved what we did in the first lesson, we did better now what we did wrong, in my case when I was told that I was not taking the prior knowledge, I was not using it with my peers either with the children, it was like… I was not using it. And now I used it. (F3, T17:30)*

> *In my case I think yes, as I said at the beginning, now I tried to do not take too many concepts because it leaves some explanations aside, let some things like random, very disconnected. I think*

I pointed the concepts much better, I focused well, I gave concrete examples and I think everyone understood. Then this... I think this was the main thing. (F3, T18:35)

The student teachers valued every part of the PA intervention, asserting each part was useful for their changing process, as it could be illustrated in this quote:

I believe that all parts of this programme were useful, because I was conducting what I did now according to the rubric. For example, 'I will do this and this' or 'I am missing that' or 'I need this, I will add this' ... 'no, I lack accuracy, then how will I get more quality in my explanation?' Then I was guiding myself by the rubric and it is good, because I feel more comfortable and able to do the things right, without being afraid of critique. I think all we did was important. (F3, T20:55)

This meaningful change of student teachers' practice was projected in their future work by six student teachers. They felt able to transfer the good practices into real teaching:

I see changes in how I do it, how to evaluate and how I'm going to perform doing a lesson and self-evaluating. (F3, T20:57)

I think that this [the transference] is going to happen even in schools where they have made the lesson planning. 'You have to do this and this.' But at the end of the day what happens inside the classroom is our decision. If you realize that what you are doing there is not good, you change it! Because you already noticed it! (F2, T13:59)

Furthermore, four participants also saw sharing the programme with other student teachers as a potential transference, especially during their first years of training.

So I think this should be done from the first year of training, subject by subject, it does not matter what it is covering in terms of content, it can be geometry, whatever, but working with the explanations it would help every subject. For example there were several things that we realized that we did not handle, and we had to prepare them for the lesson, to study and do it. So I think it's super important to have opportunities like this. (F3, T15:47)

Two teachers also mentioned the possibility of giving a wider use to the rubric.

I do not think this learning is going to be forgotten, because it's something that enriches you and strengthens you as a professional, because there are tools that the teacher does not have, and that most of us learned spontaneously only and no more, and most of us have no idea about the topics we are teaching and we teach them spontaneously only. I think in the next future it is going to be even better, because the things we learned here are going to be interwoven, for example, I could keep adding, deleting or putting other things to the rubric. (F3, T20, 67)

Actually, four student teachers suggested running PA within their future labour context in schools or placement place as another consequence of the changing practices process:

I think we should promote Peer assessment groups, so that when you receive the evaluators in the school, we would have internalized the evaluative practices. I believe that today we have a step further, if tomorrow we receive somebody going to observe one of our lessons. I believe we as a group, we can feed the rest of teacher with this. Telling them... telling them what happened here and seeing if we could work with it. We could see it like that ... Maybe it could not be done in a school with all the staff in a room, but with a couple of science teachers, 'we are seeing the teaching way issue, I could go to see you one of these days, to see what is failing, how we can improve "... Telling them a little about the work done here and obviously seeing how it could be implemented. And between us, also sharing this experience, to do not miss it ... as everyone is leaving soon. Suggest 'I could go to see you'. It is a way to keep what we already learnt here alive. (F2, T8:46, 65)

I even think this programme should be done next year, with the placement we are going to do, because there you will see yourself in this instance, for example, you come back from your placement school, and say 'today I did this and this' and show the video, and comment about the video 'this and this you have to change', and then go and fixing the weak aspects of our practice. (F3, T20:52)

I think it adds a lot in terms of our future employment skills. And this programme could continue in the labour context with our peers, to be able to improve. It could be done. If you could agree with your colleagues to assess the practice and also with self-assessment, it would improve the quality of our work, the teaching process with children, etc. (F2, T7:40)

Finally, one of the student teachers saw gaining openness to criticism as consequence of PA:

And the disposition changes also. Now I am more open to criticism. I think this is the basis of all. More open to receive them. (F2, T13, 37)

In the model presented in the next page (Figure 16), an arrow was traced from the consequences until the teaching practice. This is because it is likely that student teachers mentioned this possible transference into other contexts as another way to reflect on their teaching practice. The relationships in the model are interpretations from the researcher's view based on the data obtained from the student teachers' discourse.

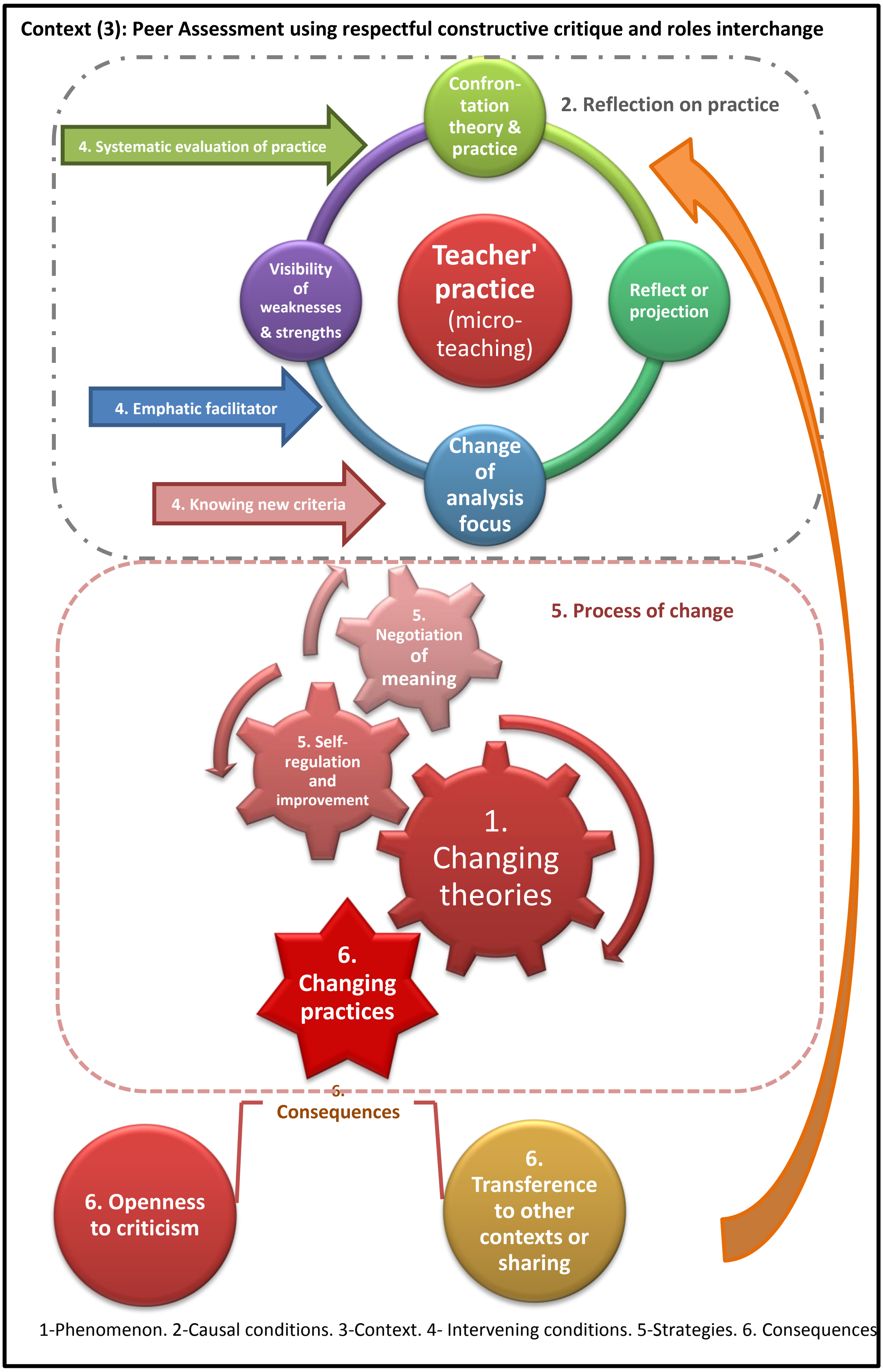


Figure 16: Selective coding: Model of changing teachers' theories and practices using PA

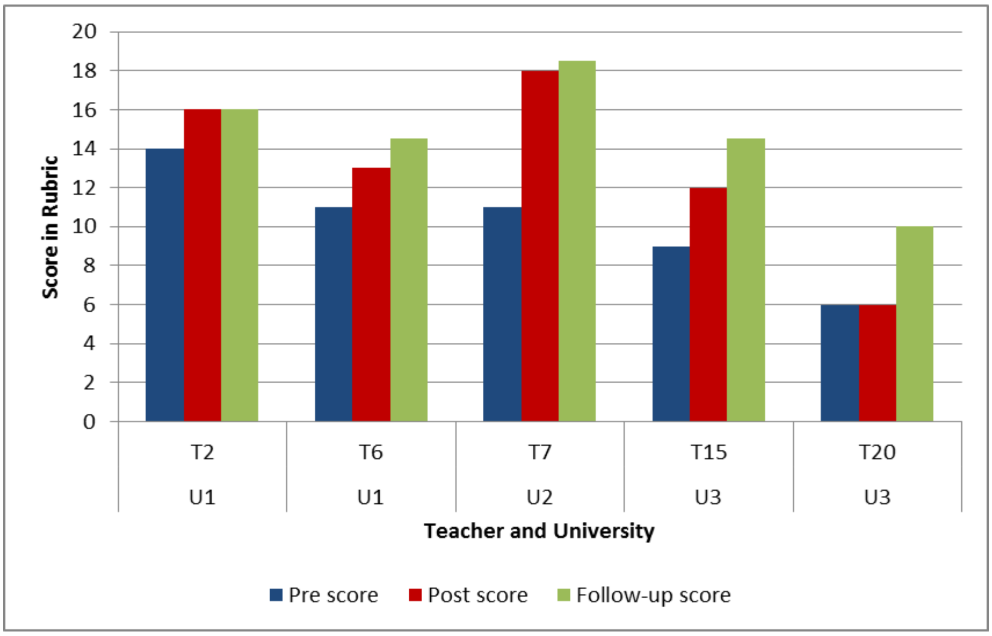
4.3. Results study 3: Follow-up of student teachers' explanations

After knowing that student teachers' practice was improved with PA programme based on the pre-post measurements, it was necessary to explore whether or not this good practice was possible to transfer into real teaching contexts and if these changes remained stable in time. Likewise, analysing the role of first real teaching experiences play in the generalization of the skill of explaining from university context into school context was considered useful for this study, in order to understand the facilitators and obstacles perceived by the teachers. For this reason, a follow-up was carried out based on case study of the participant teachers who were teaching in schools six months after the end of PA intervention, applying the rubric to a lesson video recorded and an individual semi-structured interview in their place of work. The 20 participants were contacted by the researcher through email and phone, 16 of them showed interested in this research step, but 14 of them were working. From this total, 12 of them were in teaching jobs or placement, while the two others were in educational task but with no teaching practice involved. The 12 schools in which they were working were contacted, obtaining seven positive answers and two negative. From these seven, six beginning teachers were selected accomplishing high, medium and no improvement in their skill of explaining. Cases from the three universities were taken into account. Nonetheless, in one of the schools the authorities did not authorise the observation the same day it ought to be carried out. This is the reason why the quantitative report presents five cases only but the interviews inform about the six teachers selected.

4.3.1. Transference of skill to explain into school context

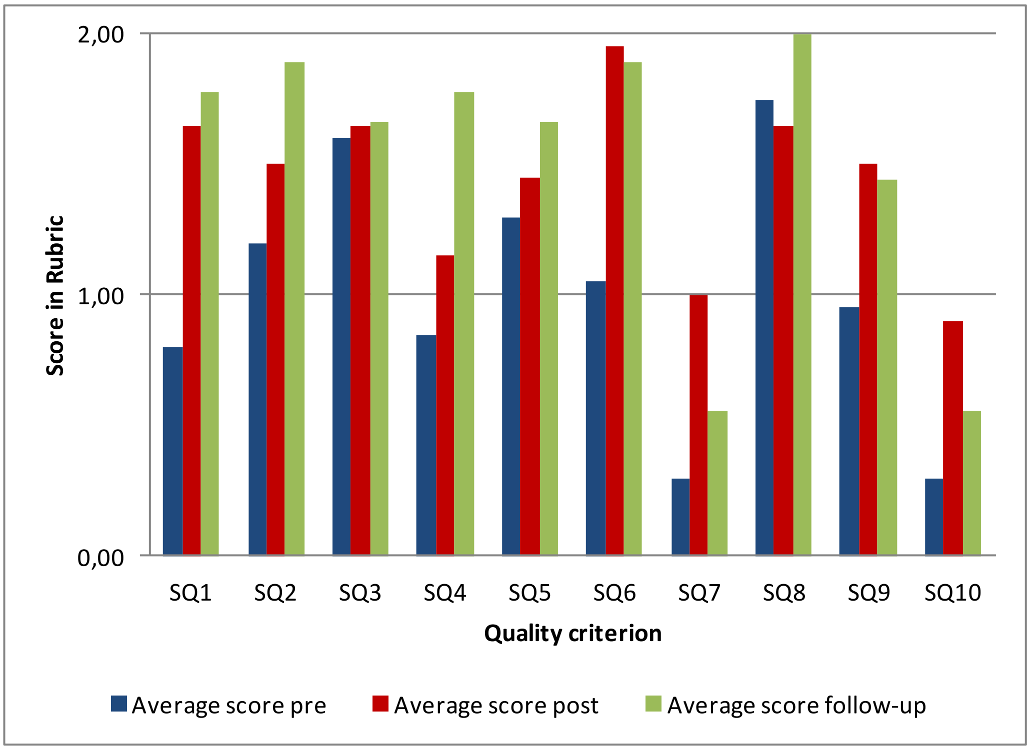
As shown in Figure 17, most of the teachers presented explanations with a better score in the follow-up study than in the pre and post measurement. Teacher number 2 (T2) only maintained his post-test score. This finding supports the idea that the improvement obtained after PA was possible to be sustained in time (at least after six months of the intervention), and that good practices to explain scientific concepts were possible to be generalised and transferred from simulated context into real teaching context. Indeed, Teacher number 20 (T20) that initially showed no improvement, in the follow-up measurement presented the highest advance.

Figure 17: Rubric score along the time by student teacher



Besides, it was interesting to observe the criteria behaviour when they were put into practice in the real teaching context. As it is shown in Figure 18, the pattern obtained with the average scores presented a higher trend in the follow-up study in most of the rubric criteria. However, the criterion SQ7 and SQ10 went down markedly, which suggests using metaphors, analogies, simulations or models and errors or common mistakes for learning were the most difficult aspects to transfer into real teaching and also to develop through PA.

Figure 18: Rubric score along the time by criterion



4.3.2. Facilitators and obstacles for the transference

As it was mentioned in the previous section, the participants of the follow-up study showed their explanatory skill maintained or potentiated when teaching in real schools. This could imply the good practices to explain scientific concepts were able to be developed during initial teacher education through PA and also generalised and transferred from the initial teacher education into real teaching contexts under certain circumstances.

To explore in the factors that could influence the transference process, six semi-structured interviews were carried out as part of the case study. The interview questions (Appendix 8.7) allowed investigating the perceived facilitators and obstacles in the skills transference process from the participants' point of view. Their ideas were organized in three main categories: facilitators, obstacles and mediators, and subcategories were created from beginning teachers' discourse emergence (see the complete map of codes in Appendix 8.11).

The analysis followed two of the three coding processes proposed by the Grounded Theory (Glaser & Strauss, 1967); open and axial. The model of the phenomenon is presented in Figure 19 (in page 150) and described in the following pages. Additional information is enclosed in brackets, showing the Interview number (I1, I2, etc.) and the paragraph from which the quote was extracted.

The transference of the skills of explaining was set up in the centre of the model as the phenomena being described. When this transference happens, the beginning teacher felt confident to explain scientific concepts in the classroom. This process could be facilitated for school conditions, and the most relevant for the beginning teachers was the presence of a tutor or guide teacher that some schools provide to new teachers. The tutor was seen as the main support in the participants' daily work, decisions and lesson plans, and it was considered by four beginning teachers as a facilitator of the transference process when perceived as a good teacher, with good content knowledge and playing a model role:

> *My guide teacher motivates me, I think she is one of the only good teachers there are in the school, because the others do not have a good academic level, she is one of the few that have it ... She is a very good teacher, she knows how to teach and she handles the content. (I2:32)*
>
> *The teacher for example, my teacher is for me a guide and a strong pillar. (I4:14)*
>
> *Once a week I seat with the coordinator to discuss lesson plans for half an hour. (I5:19)*

The second more frequent reference was the school support or flexibility. It included support to beginning teachers' lessons, and flexibility incorporate teachers' ideas, styles and interests in their way of teaching. This subcategory is shown for instance, in these quotes:

> *And the other is also the access I have to laboratory, the facilities having the laboratory has in structure and implements... Or just if you want to do something different, you have it, in a different environment, here you have the water and everything you want to work differently. Then the laboratory is a facilitator, a good facilitator. (I4:14)*
>
> *There is a lot of flexibility in this school to do the type of lesson I want. (I5:13)*
>
> *We have spoken about it with the guide teacher, and the good point is here you are listened, and they say "ok, we will do it". (I2:18)*

Another element that facilitated teachers' transference of explanatory skill into real teaching context was provided by PA intervention. Three beginning teachers valued the criteria construction process as a tool for their current practice and the critique they received in PA intervention as a way of learning and that might be internalised as self-critique.

> *I think the creation of criteria was fundamental. Because now I check it in my mind and I am going to the criterion I formulated. Because the things we saw in the university, after we do not remember it, but when you create a criteria, it is different, because you thing "let's see how I did the lesson". (I1:11)*
>
> *I think our ability to create an instrument was very important because helps us to improve our own practices. Then, from what we have created, we correct ourselves now. (I5:3)*

Actually, a beginning teacher suggested a connection between PA and the placement evaluations, as a possible development of the received programme:

> *And because several times in the placement places they go once a day or twice, but doing the explanation within the seminar context, and then another explanation in the practice place would allow us to evaluate and compare reality with what we can do during the peer assessment seminar. (I4:21)*

A few factors that may work as mediators were identified by the researcher in the beginning teachers' discourse. The mediational function was interpreted because these factors might make stronger or weaker the effect of the facilitators or obstacles on the transference of the explaining skills. It means, depending on their presence or absence they could become a facilitator or an obstacle. Indeed, some beginning teachers mentioned as one or another independently. Tree main mediators were identified:

Firstly, the quality of classroom climate the beginning teacher was able to generate was mentioned. Five beginning teachers identified it as a potent mediator that could facilitate or, in most of the cases, it might hinder explaining in the way they wanted.

> *But only when I generate this good classroom climate, I can explain in the minutes that I am achieving of real lesson. (I1, 13)*
>
> *I believe that children sometimes they are not staying quiet, and then they do not listen. So if they do not listen, you lose the connective thread of the explanation, because there are people making noise, disturbing, and there are others pupils concentrated, then you think, 'Why they are not listening to me, maybe they do not care." And I think that makes teaching more difficult sometimes. (I2, 26)*
>
> *[It is very important] the issue of children and their behaviour in classrooms. (I3, 16)*
>
> *There is also the factor of classroom climate. Here the pupils are not so messy, but it is not the ideal context. There are always pupils that disturb the class, which are putting others pupils off. (I5, 7)*

Secondly, the school resources were mentioned by four beginning teachers, as potential facilitator of the generalization of their skill or being an obstacle for it, as shown below:

> *Also, the resources are given to you at this school, although they are basic. We have computers, we have materials, have back garden, theatre for a play, etc. (I2, 17)*
>
> *It is a real need, but if the school does not take it ... I think the school knows [the resource problem] but it is limited to the resource. I believe that. (I1, 52)*
>
> *Here there is very little material to create work sheets, all that is printing I need to do it, I pay for it. So in that sense I would like to have more support, in order to do more exercise sheets, more explanatory sheets, having that resource ... I just want them to print it and distribute it to my two grades, fourth and fifth grade... that would facilitate a lot my work. (I3, 22)*
>
> *Also, here I have access to computers, internet, printing, photocopying and nothing is charged to children or teachers. (I4, 14)*

Finally, the teachers' content knowledge and knowledge of pupils was considered relevant by four beginning teachers, because as they were recently joining a school, their knowledge of pupils was little. Also they felt not confident about their content knowledge. They thought that having more knowledge of pupils would make easier transferring the good explaining practices, i.e. incorporating pupils' interest in the explanation.

> *The first thing is as I am newcomer here; I do not have prior knowledge about the children. (I3, 10)*

> *Regarding the explanations, they are always more related with the knowledge the pupils already have than with what I want to teach. (I1, 27)*
>
> *I advance in certain areas, depending on the subject matter, I mean depending on how I handle the content. (I2, 12)*
>
> *In my own skills, I would like to structure more ... not the way to do the lesson because I can keep it, but the conceptual part, I would like to have it more incorporated, to be able to give it less time to do that. (I3, 20)*
>
> *[The difficulties] are mostly due to insufficient training I had, it is an obstacle in my case because it makes me doubt even more than I doubted before. And my weaknesses are in content and methodology. (I4, 17)*

Some of the factors described above were felt by the beginning teachers as consequence of their initial teacher education. In their words, their education as teachers was weak and disconnected from real teaching contexts. This might make them feel less confident at the moment of explaining to real pupils, then, it was understood as a source of difficulties.

> *But I think that lessons we received in the university were planned considering that we will have an ideal class, where students' skills are high, where the classroom climate is good, where we are not considering the problems that students are exposed. (I5, 11)*

Considering the points that beginning teachers mentioned directly as obstacles, it was possible to identify four main factors. The two first of them are associated to pupils: their lack of science knowledge and lack of participation during the lessons. These were mentioned by four teachers, as shown in the following quotes.

> *Well, in this case I could easily explain and explain generating a monologue. But when you ask questions to the students and you make students participate, you notice here students do not participate when I ask them. (I1, 29)*
>
> *The pupils have the concepts still very basic... I have to start levelling out to be able to teach. (I3, 14)*
>
> *The pupils had problems with previous teachers, and then they have some content deficiencies. (I5, 7)*

Another element that markedly appeared in teachers' discourse was the little time they had to plan the lessons, prepare the materials, evaluate students' learning and reflect about their own teaching. This was a very frequent factor mentioned, nine times by two of the teachers:

I would like to have more time to prepare lessons, because time is a crucial determining factor. I dedicate the weekend to do it, in between that I have to have time for family life, and now for example I am taking paper work to be done while I teach the other lesson. (I3, 16)

I stayed working at home, and I try to mark the tests always the next week the student took it, and it is a marathon job marking 44 tests, checking the grading scale, putting the marks in the lesson's book and bringing them back. The marking process is done immediately. (I1, 46)

As I have little time for preparing the lessons, I would like to distribute less time to it and being able to apply only what I already have structured. I would like to do that. (I3, 20)

I believe that more than the lack of time for planning, the problem is the time for teaching reflection. When I am supposed to do my teaching reflection? On my pillow? (I1, 57)

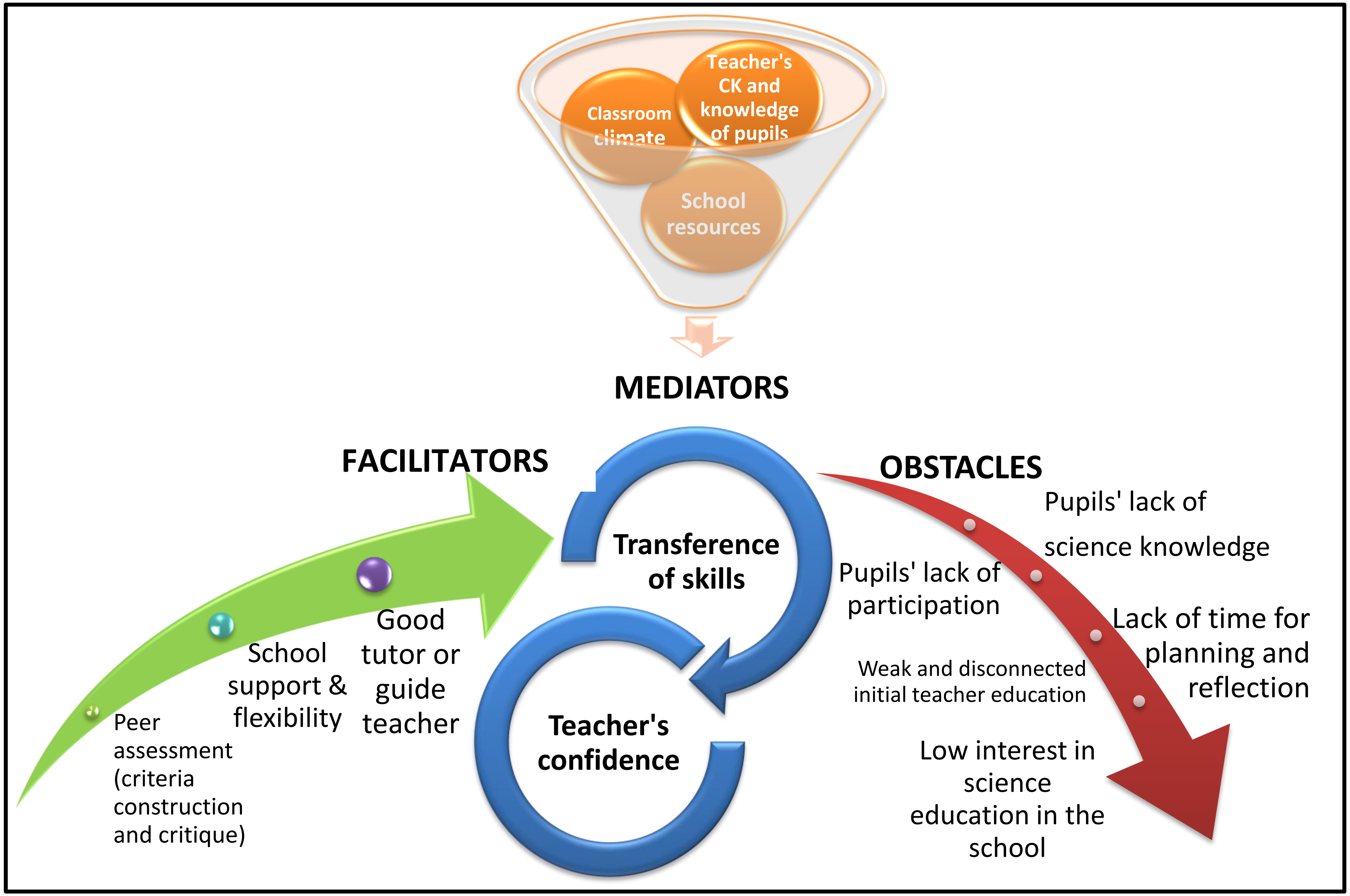
Finally, teachers mentioned that in some schools trying to deal with low interest in science education was a difficulty. Two beginning teachers perceived this as an obstacle to get pupils' motivation for learning science.

In this school, science as an everyday life thing does not exist. (I1, 33)

Then, when you try to focus them and interest them to learn what you are showing… Even now that they are only 5 or 6 pupils who are really interested in learning sciences, because they are more interested in learning music or other things. They are also important areas to teach, but getting motivation for science study is difficult. (I4, 8)

How these factors act together putting pressure on the transference of skills is graphically presented in Figure 19. The relations between the components are based on the researcher's understanding of the phenomenon interpreting the data gathered from the beginning teachers. In the model, the font size shows how frequent each aspect was mentioned by the participants. The components of the model are important to take into account to coordinate schools and universities in helping beginning teachers to transfer the skills they may have developed in the initial teacher education into their labour context.

Figure 19: Model of factors affecting the transference of the skill of explaining from university into school context



5. Discussion

This chapter summarises the main results obtained, leading to a discussion of the methodology used, suggestions and alternative reasons for the results. The limitations of this research and descriptions of how those were addressed are included, discussing further the generalizability and significance of results. This is followed by linking back to the results and the literature review, based on empirical facts and interpretations from the researcher's analysis. The chapter ends with the main contribution to theory on peer assessment (PA) in the context of initial teacher education (ITE) and about the quality of teachers' explanations.

5.1. Summary of results

The research conducted was directed by three studies. The first study consisted of the construction of an instrument to evaluate the current quality of science pre-service teachers at the end of their final year of university training. The result was a rubric with ten assessment criteria and three achievement levels. It was reliable using Cronbach's Alpha indicator (α = .77, n=10) and valid according to an expert panel. Its reliability enables the assessment of the skill of explaining based on ten elements identified in literature.

The second study introduced PA intervention to three groups of student teachers, measuring the quality of their explanations at the beginning and at the end of it. Results showed pre-service teachers significantly improved their explanations of scientific concepts in most of the elements identified by the rubric ($p<.05$). Here also the teachers' conceptions were measured and compared with a control group. Qualitative analysis indicated how teachers transformed their conceptions about the quality of explanations from general pedagogical knowledge into pedagogical content knowledge during PA intervention. The control groups' conceptions remained stable in this time. These changes were analysed from the teachers' and researcher's point of view to indicate how the process occurred and to what extent PA had a role on it. The third study was a case study follow-up looking for the generalisation and transference of the skills of explaining elements into science classroom. Results indicated that eight of the ten elements were successfully transferred. The clusters: using analogies, metaphors, simulations or models; and using errors or common mistakes in the understanding of the concept as a learning opportunity were not generalised. Interviews conducted indicated from the beginning teachers' perspective which factors facilitated or made difficult the transference of good practice to explain into a real teaching context.

5.2. Limitations of methodology

The results described in the previous section regarding teachers' conceptions, theories and explanations were obtained using the research methodology detailed in Chapter 3. To summarise, a ten-session PA based on microteaching was conducted in three Chilean universities differing in the quantity of science courses they offered to their students. In this research, the 20 student teachers who voluntary attended the ten-session PA programme composed the three groups. Also, 18 teachers were divided in three control groups (one in each university) and completed the initial and final PA questionnaire. Student teachers' thoughts were gathered in all groups through this instrument. In the experimental groups the PA questionnaire was complemented with feedback sessions, focus groups and interviews as data collection techniques. The qualitative data collected in this study was analysed using content analysis, positioning analysis and constant comparative analysis. Otherwise, the quality of teachers' explanations in experimental groups was measured at pre and post via video-recorded microteaching episodes using observational analysis. The same analysis was used in the follow-up study via recorded teaching lessons in schools. A quantitative instrument was created to evaluate student teachers' explanations in practice. Inter-rater reliability was calculated on 5% of all qualitative data and all the videos were rated by two researchers in a blind marking process.

The three studies that comprised this research have offered an understanding about the development of the skill to explain scientific concepts using PA. In this sense, the methodology was adequate in general terms to achieve the research objectives. Nevertheless, as a consequence of implementing action research based on professional development workshops, field work constraints, sample reduction and decisions taken by the researcher and participants, some limitations were encountered. These limitations are critically analysed below, to be considered for further research and interpreting the results of this study.

First of all, as it was mentioned in section 3.5.1, it was not possible to have random selection of the student teachers during the sampling process. This was due to requirements from the participant universities, which preferred running this project voluntarily for their students. This was contrary to the initial agreement, but when the field work of this research started there was a significant student demonstration and strikes that did not permit conducting the normal university lessons as planned. In this context, the universities gave priority to completing the students' compulsory curricular activities and any extra-curricular activity (as the PA project) was settled in a voluntary character. As a consequence, the research sample was considerably reduced in size regarding the initial expectations. Besides, it was not possible to have random assignation to groups or random selection of the cases considered for the analysis. Indeed, all the participant teachers were considered in the analysis and for some analysis the three experimental groups were taken as a whole group to reach a reasonable sample size. Bearing this condition in mind, self-selection variables could have affected the decision to be part of the programme. The participants of this research represented a selected group and not the totality of pre-service science teachers of the universities. Nevertheless, possible differences between the three experimental groups were explored by running Levene's test for variance normality to analyse the distribution of the variables. Also, similarity between control and experimental groups in each university was assured by quantitative and qualitative comparisons of their outputs in PA questionnaires before and after the intervention. The control and experimental groups did not show marked differences at the beginning of the intervention, although there were differences at the end of it. These differences were attributed to the PA intervention.

Secondly, and associated with the previous problem, during the data gathering it was not possible to have a comparison group available to experience delivery and recording of conceptual explanations. The control groups were available only for completing the PA questionnaires. It was a potential weakness at the moment of interpreting the advance of teachers' skills as produced by the PA intervention. However, this point was intentionally explored in the focus group questions, investigating if the changes perceived by the teachers were spontaneously connected with PA from their perspectives. In fact, the student teachers mentioned that all the improvements they had experienced were caused by different components of the PA intervention. In spite of this, having a comparison group to conduct the analysis of teachers' explanations is considered as an ideal scenario to replicate this research.

Thirdly, in order to respect ethical considerations when researching with people, the student teachers were not forced to stay during all the sessions of the programme. However, only teachers with an established minimum of sessions were considered for the analysis, which also importantly reduced the sample size. This constituted a limitation for running other statistical analysis such as factor analysis or multi-level modelling to explore the rubric characteristics in the instrument validation phase. If similar research is conducted in the future, it is recommended to establish a better incentive to encourage participants to stay until the end of the research period, and having further mechanisms to maintain the sample during the time.

Furthermore, the teacher knowledge variable was operationalized according to the number of science courses that each university offered to their students. This might be an issue because it is recognised that initial teacher education is not the only source of scientific knowledge for students. As Arzi and White (2008) have mentioned, knowledge that teachers use in classroom teaching does not start and end in university, there are other factors that can strongly influence it such as their school learning as students or their life out of school. Further research might include a valid standardised knowledge test for student teachers, in order to have a more accurate measurement of this variable and consequent sample stratification.

Finally, as the content of the rubric was largely developed with respect to the literature review, further studies both on the content aspects and the statistical analysis should be carried out in the future. The rubric needs to be customised to be appropriate to different cultures of teacher education, because it is known that each university might have different views about teaching how to teach science. As this instrument fits better with teacher education programmes that apply constructivist theory of learning to their programme design, this possible limitation needs to be addressed if the instrument is applied to different teacher education programmes' views. For instance, contextual adaptations or checking the rubric items by expert teachers from the context where it would be applied is recommended. Likewise, the rubric used in this study contains criteria that apply specifically but not only to science teaching, then, adaptations to other subjects of teaching education is proposed. The points discussed above indicate a need to treat the findings of this research with care. Also, consideration of alternative explanations for the results obtained was required. Thinking about other possible explanations for the results obtained in this research, recording student teachers' practice could itself improve their teaching because of their own observation and analysis. Similarly, student teachers could have a positive advance due to the specific attention paid to their individual case which is not usual in large classes such as in Chile.

Additionally, it is important to mention the possible effect of the teacher students knowing they were participating in a study. Although the researcher did not tell them what the specific hypothesis or expected results were, this was easily identifiable as peer feedback was given in order to help them to improve their teaching, and this might have generated student teachers' commitment to the research objectives thereby enhancing positive results.

5.3. Generalization and significance of results

Considering the conditions of the research sample, it is important to question about the generalizability of its results. The universities that participated in this study were representative of average institutions which prepare teachers in Chile; they received students with similar prior knowledge (medium according to the National test for University Selection) and similar economic income level (low-medium, according to national parameters in higher education). These institutions represented the most typical teacher education programmes in Latin America (see section 2.3.2 for the definition). In this sense, the results of this research are considered as highly generalizable to other teacher education programmes in Chile and also in the Latin American region. However, the results need to be interpreted with caution if applied to developed countries with more selective entrance requirements to teacher careers, or programmes that receive students with a high level of science knowledge (e.g. Western Europe) as students' characteristic may vary.

Otherwise, due to the voluntary character of the sample in this research, it is not possible to assure the results would represent all of the teacher students because hidden variables may have affected the self-selection. Nevertheless, as in the three university groups, high, medium and low quality of performance were observed before PA intervention, therefore perhaps the groups were representative of the complete scope of the skill studied.

Regarding the question if good practices obtained after PA would last over time; the follow-up study provided clues to understand the transference of skill process. In the group of beginning teachers studied, the trend was maintaining or potentiating the advance of the skills they acquired as pre-service teachers during the PA intervention in the schools, considering a medium time after the intervention (6 months). This could lead to a positive projection about the maintenance of these results. However, it is necessary to consider the first years of teaching practice are often the most important to model practice, and in this period new teachers receive different influences in their teaching from colleagues and superiors (Day et al., 2007; Schubert, 1992). Thus, it would be essential to keep monitoring and giving teachers the space for critical reflection of their own practices to avoid the disappearance of the positive effects shown in this research. Indeed, most of the beginning teachers who were followed-up in this research commented about the relevance of PA between colleagues in their place of work in order to keep improving their practices and solve tensions encountered in teaching in different contexts, which is an interesting suggestion coming from the research participants.

In terms of the interpretation of significance, the results are considered educationally significant because there was a high effect size of PA according to Cohen (1988). These results were not insubstantial as they could be due to the small sample size. There were statistically significant differences in how pre-service teachers explained concepts by comparing their performances at the beginning and at the end of PA intervention. Student teachers maintained this positive advance when teaching in real classrooms and even in some cases it improved further. Assuming these results were at least in appearance and from participants' perspective facilitated by the PA intervention, it is valuable to link them with other ideas coming from different researchers or authors. This is described in the next section.

5.4. Discussion based on empirical facts

In this section, the literature related to the research results was considered, making connections between other researchers' findings or authors' perspectives, which are described as follows.

Microteaching situations in this study elicited student teachers' beliefs and assumptions about teaching science, which in their words were not addressed in any workshop before, as can be visualised in this quote:

> *I do not think this learning is going to be forgotten, because it's something that enriches you and strengthens you as a professional, because here we had tools that the teacher often does not have and that most of us learned spontaneously only and no more... Most of us have no idea about the themes we are teaching and we teach them spontaneously only.* (F3, T20:67)

This fact is similar to the idea presented by l'Anson et al. (2003), where microteaching was introduced in their initial teacher education (ITE) programme to enable student teachers to become aware of the nature of their inscribed values, attitudes and assumptions about learning previously internalised. In this study, these values, attitudes and assumptions informed their practice within the microteaching space, as similar outcome as in the present study.

Also, the reflection and projection mechanisms that from the researchers' point of view in this study were underpinning why PA in ITE worked, can be explained by the statement of l'Anson et al. (2003). These researchers asserted that, in general, most student teachers in their ITE programmes did not have a repertory of teaching alternatives, and that they have not anticipated things which are not present in their experience yet. The opportunities for multiple viewing of their teaching may influence the future development of their teaching repertories. In the current research it was possible to observe that teacher students projected their own field of decisions and repertories in their peer's teaching practice. Consequently, student teachers could see and evaluate different methods of explaining science concepts that they had not tried before. Thus, PA based on microteaching episodes widened the repertoire of alternatives that l'Anson et al. (2003) indicated, but in the current time. This might accelerate the progress of teachers gaining experience because they did not have to wait until a future real teaching period to explore different teaching as proposed earlier.

In this research, from the participants' point of view their perceived improvements were attributed to the existence of respectful and formative critique, confrontation of the practice with their own theories about teaching science and the change of analysis focus they developed thanks to the instrument construction to evaluate peers' practice. These findings are complementary to Hume's perspective (2012), who reported using role-playing in simulated primary science teaching facilitated teachers' development of science content knowledge (CK) and growing awareness of pedagogical content knowledge (PCK), because the role-playing allowed them to reflect about issues and problems that they may face as novice science teachers. In the present research, the former perspective was widened.

In general terms although the advance was different in the three universities in the present research, the three groups developed their initial levels of the skill of explaining scientific concepts. Actually, though a better performance in student teachers with a high content knowledge level was expected, results showed that these elements were not necessarily related. The student teachers with a balanced curriculum between pedagogy courses and science courses (University 2) performed better at the beginning of PA and at the end of it in than the student teachers in the lowest level of science knowledge (University 3) and the ones with a highest level of science knowledge (University 1). This finding was concordant with results reported by Lloyd et al. (1998), who compared two groups of teachers with similar levels of subject content knowledge and pedagogic ability, and found in neither group was any clear connection between science knowledge and the ability to teach that knowledge. However, this result contradicts the idea of Shulman (1986), who implied teachers cannot craft PCK and explanations until they are content experts and also expert pedagogues, which happens when they have several years teaching the same subject.

In this sense, this research has shown that the roots of PCK can be developed during initial teacher education, which resonates more with other researchers' points of views who have intended to challenge Shulmans' postulate such as van Driel et al. (2002), Tsangaridou (2002) or Geddis (1993). Nonetheless, the practical component needed for the development of PCK established by Shulman (1986) was in the same line of this research results: the student teachers settled in the centre of their perceived changes the teaching practice that microteaching allowed them. The variation here with the conceptualization of Shulman (1986) is that teaching practice could be a useful tool even in simulated contexts in early stages of teachers' experience.

If explaining is considered to be a key competence of science teaching as positioned by Geelan (2012), the contribution of this research to develop this competence using PA is more important, because it demonstrated the development was sustainable. In this case, the explaining scientific concepts competence was recognised, confronted, developed and put into practice. In this regard, Barnett and Hodson (2001) have stated that developing competences in science teachers implies recognition of when to invoke and how to apply contextual knowledge, recognising how generated strategies and contextual knowledge interact. Thus, experts have more accessible and usable knowledge than novices. In the present research this was observable when the student teachers described as having the criteria in mind when analysing their own teaching:

> *I think the creation of criteria was fundamental. Because now I check it in my mind and I am going to the criterion I formulated. Because the things we saw in the university, after we do not remember it, but when you create criteria, it is different, because you think "let's see how I did the lesson". (I1:11)*

Participant teachers in this research advanced in their grades of organization and approachability to professional knowledge. The capacity of change is a significant aspect of reflective practice, and the process of reflecting-changing involves making thoughts about practice explicitly (l'Anson et al., 2003). In the current research student teachers made explicit their thoughts about quality in science teaching and specifically in explanations through the PA process. This was a justified assessment, it means, the facilitator asked them to indicate where their ideas were coming from to judge as better or worse the peer's performance. In the work from l'Anson et al. (2003) it was the opportunity to revisit a microteaching episode through the eyes of peer, tutor or fellow teacher and the same student teacher which created the potentially powerful event leading to teachers' reflection. In the present research it was proven that doing this from the peers' eyes only is powerful enough to facilitate changes in teachers' thoughts and practices, then, its efficiency can be recognized.

An important gap in the literature review was found regarding science teacher explanations, specifically in the quality criteria that could make an explanation of better or worse quality. In this sense, although some authors presented characteristics such as accommodation to the audience (Leite et al., 2007), understandable (Wragg & Brown, 2001), containing examples (Eder, 2005) or avoiding tautology (Faye, 2009; Geelan, 2012), any author gave a light about how to use those characteristics as an assessment criteria. It means transforming the characteristics into observable and evaluable quality criteria was not done. Additionally, empirical studies in science teaching that have judged the quality of general science teaching such as Goodwin (1995), Leite et al. (2007) or in science teacher education such as Lederman and Gess-Newsome (1989) or Ginns and Watters (1999), did not indicate how the quality of science teaching devices was measured, how the changes were established or in some cases, what the criteria to establish the possible improvement were. In this sense, the creation of a rubric to evaluate the quality of science teacher explanations, with observable criteria in different achievement levels, useful to diagnose and intervene in ITE or in teacher continuous development is valuable as an original contribution of the present research.

Another relevant gap in the literature review was about the generalisation or transference of skills gained in PA to broader or different contexts. Usually the studies in PA do not include follow-up and it has been stated as a need for future research (Sluijsmans, Brand-Gruwel, van Merriënboer, et al., 2002). This was important because according to Barnett and Hodson (2001), to function satisfactorily in science teaching and develop the unconscious and seemingly “automatic” quality of teachers’ work that makes them go from novice to experts, teachers must be able to generalise some aspects of knowledge and skills to new situations. Furthermore, Pauline (1993) stated that the main problem of microteaching as a learning tool is its differences from classroom settings, because the skills gained there would be difficult to transfer. However, in the present research this gap was filled investigating in the transferability of the skills to explain scientific concepts into real teaching. Results of the follow-up study showed that participant teachers not only generalised the skills acquired during PA of microteaching episodes to other teacher contexts, but also maintained eight of the ten practical aspects to explain scientific concepts. This denotes PA can be successfully used in ITE to improve aspects of science pre-service teachers’ explanations such as clarity, coherence and cohesion, sequence, sufficiency, connection with students’ experiences, use of concrete representations and use of non-verbal emphasis as shown in this research. However, in this study the improvement of using metaphors, analogies and models or common mistakes as sources of learning was not high, and those elements were not well transferred. It might be because these elements require a more mature content knowledge development, which has been argued by Davis and Petish (2005), stating that content knowledge plays a crucial role on teachers’ instructional representations, and a well-integrated, principled, and scientifically accurate understanding of science would lead to more pedagogically appropriate teaching. In this line of discussion, the present research contributed to clarify which aspects can be transferred and which others present more difficulties in a group of beginning science teachers.

Finally, there was lack of studies with experimental designs that could affirm PA or microteaching usage in ITE had positive results as a consequence of these strategies (Kollar & Fischer, 2010; van Zundert, Sluijsmans, & van Merriënboer, 2010). In most cases it was due to research which did not consider a comparison group and random selection to establish the causal relation. This was intended in the original design of the present research. However, as it was detailed in the limitations of the methodology of this study, it was also not possible to have perfect experimental conditions. Thus, the attribution of results to PA intervention was not found from quantitative data but it was done by qualitative data analysis. From participants' perspective, their perceived changes were spontaneously connected with PA intervention during the focus groups. Actually, they mentioned all the improvements they had were the result of creating the quality criteria, performing and receiving feedback.

5.5. Discussion based on interpretations

Going beyond the data, it is possible to make interpretations regarding several of the results and also with theories or previous work that might give complementary perspectives to interpret the present research findings

In this study, teachers constructed collectively an instrument to assess peers' explanations which likely led to negotiate the meaning of quality teaching. In terms of l'Anson et al. (2003) becoming a reflective practitioner in the teaching field implies an on-going process, consisting of negotiating thresholds rather than a particular outcome. They asserted individual negotiation of meaning was possible due to engagement with wider discursive perspectives that imply their individual repositioning. In their work the student teachers' opportunity for dialogue with tutors, fellow teachers and peers encouraged student teachers to participate and negotiate before they joined a community of practice teaching in schools. Otherwise, Barnett and Hodson (2001) highlighted that feelings, attitudes and personal aspirations of student teachers interact with their process of cognitive restructuring. However, these two views have been focused mostly in the individual level of restructuring. The current research complemented it with the role of PA in the restructuration of student teachers' theories based on peers as a source of social construction of meaning, then it goes a step further the previous mentioned authors.

Furthermore, in this research, student teachers associated the transference of the skills they considered to have gained after PA to the confidence they feel about teaching science in classrooms. This can be interpreted because criteria development allowed teachers to have a more easily justifiable practical knowledge. It means they could access to the criteria they created to justify why they made particular decisions or how they evaluated their own performance. This is comparable to the view of Barnett and Hodson (2001), who indicated teachers' personal practical knowledge has two functions: to provide teachers with a sense of personal control -knowing what they are doing and the confidence to feel they can do it- and provide them validation as a teacher. Probably this was why teachers in this study associated their skills transference with their confidence when teaching scientific concepts at school.

The role of PA in ITE can be seen from the problem of developing PCK, which is a very contextual and situated knowledge as well as being necessary for effective teaching (Onslow et al., 1992), when there is no context of teaching practice develop it. In the present research, having an opportunity to teach but also to observe peers' teaching in several episodes, science concepts and teaching methodologies makes sense as an opportunity to develop PCK, as shown in the teachers' conceptions analysis based on the feedback sessions and assessment questionnaires. Barnett and Hodson (2001) have indicated that teachers' knowledge is rooted in details of particular classroom experiences, especially in those that present difficulties, because is in these circumstances that personal theories are put into action. Thus, the opportunities PA intervention offered to student teachers were valued by them regarding this aspect.

It is also interesting to understand why there were two criteria which were more difficult to transfer than the other eight that the rubric considered. These criteria were: using analogies, metaphors, models or simulations; and illustrating error or common mistakes as an opportunity for learning. On the one hand, another research with pre-service science teachers found that although they certainly posed the ability to create and utilize analogies, they were unlikely to use this strategy when teaching unfamiliar concepts (James & Scharmann, 2007). This might be a useful explanation considering student teachers in the present research prepared their microteaching episodes and choose the concept they wanted to explain (even though they were encouraged to explain a concept which was perceived as difficult), but during the follow-up study the participants were teaching the content matching the national curriculum for that period of the academic year. Thus, these concepts may be less familiar for the beginning teachers and probably the observed lesson was the first time they were teaching them.

On the other hand, metaphors, analogies, simulations or models are usually understood as devices that require a higher level of knowledge and thinking skills (Ogborn & Martins, 1996). Besides, it has been indicated when teachers face real teaching they often reduce their expectations of children's performance because they find pupils have less knowledge about science and a lower level of thinking than they expected (Bryan & Abell, 1999). Indeed, student teachers in the present research mentioned pupils' lack of science knowledge as one of the important obstacles for them to teach science in the way they would like to do it. Another possible explanation for the lack of usage of analogies, metaphors, simulations and models observed in the follow-up study regards the noticeable focus beginning teachers had controlling the class more than in the teaching strategies they applied. This could be assumed because for some beginning teachers the consequence of the first experiences of teaching make them reject their student-centered views and adopting a more conservative controlling stance (Geddis, 1993). Beginning teachers in the current research showed an important preoccupation about their classroom management, as it was also found in the work from Woolfolk-Hoy and Murphy (2001). Actually, in some more extreme cases reported in another research when pre-service science teachers start working in schools they may have a "reality shock", feeling extremely anxious about their performance (Bryan & Abell, 1999).

Related with the previous ideas, the participants in this research showed concern about the quality of their explanation when they were observed in schools, because of the high levels of noise in their classrooms and that there were several pupils not concentrated in the explanation. From the point of view of the researcher, even though the participants were not very successful in classroom management, this did not imply that the quality of their explanations was poor according to the criteria presented in the rubric. Actually they seemed to self-evaluate their explanations focused on the classroom environment where they were explaining more than in the explanation itself. Thus, the development of teachers' skill to identify good practice mentioned by Sonmez and Can (2010) makes more sense. The teachers need to be able to analyse and isolate the components of their classroom settings to identify causes and effects and try out improvements. If not, they would continue making wrong judgements about their performance which can affect their sense of professionalism and decision making process. Also, it is valuable to cite here an idea from Treagust and Harrison (1999) who declared "teachers who are conscious of the constraining influence of the science content, the educational context, the students and their own teaching and content knowledge limitations are more likely to recognize the challenge posed by classroom explanations. Indeed, teachers who purposefully reframe some of their explanations in light of these factors will likely enhance the quality of their classroom interactions" (pp.40-41).

Further, in terms of the generalization of the results to other contexts, from pre-service teachers' views there was a strong emphasis on the possibility of transferring PA to their workplace context. This could be interpreted as the need of developing a more professional identity in the schools, which has been addressed by Boud (1999). This author stated that teachers' academic development could be potentiated as a local practice based on process of peer learning in the workplace. In this line of development, other researchers have stressed teachers can become more professional teachers in an atmosphere where they can interact with other skilful teachers in direct ways, or with mentors because they begin to understand teaching strategies in action and constructing their own style (Borman et al., 2009; Darling-Hammond, Wise, & Klein, 1995). Consistent with this logic, participant teachers' emphasis given in this research takes more relevance.

In addition, it is interesting to observe the development of PCK in student teachers of this study on the perspective of the conceptualization of *teacher lore* which has been described by Schubert (1992) as the powerful oral tradition by which ideas, perspectives, insights, images of teaching and everyday strategies are passed on to new teachers. As the teachers from this group have already more advanced knowledge to justify their decisions, it is likely their relationship with teacher *lore* would be more active or richer and not only received as a one-way transmission. Thus, the student teachers would be more able to negotiate meanings in the near future as consequence of the PA processes. Otherwise, an important gap the in literature review about PA was the understanding of how it works in assessment of performances such as in teaching during microteaching episodes. In this sense, the researcher interpreted from teachers' comments and the position they adopted in their discourse when giving feedback to their peers, that two mechanisms were having a role in the of PA as a facilitator of teachers' change process: the projection and reflection as described in section 4.2.5. This finding enriches the understanding of the underpinning principles that could make peer feedback and assessment a formative tool.

To summarise, one of the missing points in the processes of cognitive restructuration in terms of teachers' conceptions was the role of peers in the social construction of meaning. In this argument, although its importance was stated, the available studies in PA or microteaching did not offer a conceptualization of the restructuring process itself or what the roles of PA elements were. The present research proposed a model based on reflection and projection mechanisms. Likewise, the research results added new elements to the understanding of the teaching experiences as a key component of PCK development, showing it can be triggered even through simulated teaching contexts when their conditions are carefully prepared. Moreover, this piece of work contributed to confirm that metaphors, analogies, simulations or models and using mistakes as a learning opportunity are difficult to transfer from initial teacher education to science classrooms. Although the first group of teaching devices have been widely studied, this research corroborated it expounding the reasons why this might happen.

Finally, although it has been stated that the skill to observe, analyse and evaluate practice is needed in pre-service teachers analysis to make the best when using videos as a learning strategy, little attention has been paid to the role of their abilities to recognize strength and weaknesses in their own teaching (Sonmez & Can, 2010). The present research contributed to show how relevant that skill in pre-service and beginning teachers is, and that it might be helped with the creation of assessment criteria to assess not only their peers' but also their own strengths and weaknesses during science teaching. This ability can be trained in simulated contexts such as using videoed microteaching episodes. Besides interpreting the findings of this research from the literature review, it was crucial to determine the contribution of this study to theory, which is presented in detail in the next pages.

5.6. Broader implications for theory

In terms of the implications of this research for theory, PA in ITE has been positioned as quite a new area of research in general (Kilic & Cakan, 2007), and most of the studies have not been focused on science teacher education as was argued in theoretical discussions presented in Chapter 2. The few studies found on this subject were focused on student teachers' products or performance marks between peers and instructors, looking for reliability (Kilic & Cakan, 2007; Tsai et al., 2002), but not stressing the formative power of PA itself in science teacher education as in the present research. Here, it was observed PA can facilitate changes in practice of pre-service science teachers, which was yet unexplored using this methodology. In this sense, the present study can illuminate an area that was not characterised before, increasing the range of studies available in the usages of PA which is valuable as an original contribution.

Furthermore, the projection and reflection mechanisms that were found in this study having a role as a facilitator of teachers' change enriched the understanding of the underpinning principles that could make peer feedback and PA a formative tool. In this area, it was already mentioned in Chapter 2 that students might learn better from others in a similar learning level because they trigger their peers' learning (Vygotsky, 1986). This research contributed to the understanding more specifically of what happens in the context of PA and peer feedback in teacher education as mediational processes. A graphic conceptualization and an explanation using the analogy of a mirror are presented in Figure 20 [1] in next page. This analogy was inspired by the work of l'Anson et al. (2003) but differs consistently from it.

Moreover, these described mechanisms could influence further development of theories not only about why PA works but also why PA is usually reported as difficult to carry out with friends or close people, as a study previously suggested (Woolhouse, 1999). The findings in this research would not suggest keeping anonymity in PA as it has been mentioned by Woolhouse (1999) or Bostock (2000), because the subjectivity and social interaction between individuals with similar experiences and theories are in the origin of these mechanisms. Probably anonymity in PA would be useful in terms of reliability based on

[1] Authorization to use the prism image was obtained from Graciela Lobos González, Physics teacher from Pumahue School in Temuco, Chile (in Appendix 8.12). The original image was duplicated and modified to give the mirror shape and text was added to illustrate peer assessment components.

quantitative marks, but it would be less beneficial in a qualitative formative feedback process because the individualities create the space where projection and reflection mechanisms can act.

As shown in Figure 20, the assessors projected their own decision making on the peer's performance and the assessee reflected what the assessors would do in a similar teaching situation. A model beyond the data is presented in Figure 2. Each line represents an element, organized from letter (a) to (h), as described below. (a) Teacher 1 (T1) is required to create a microteaching episode and (b) creates a scientific concept explanation. T1's performance is represented by the T1 prism. (c) As in a mirror, reflection mechanism operates: T1's practice, decisions and mistakes reflect what other teachers (T2, T3) would do. (d) Teachers in the assessor role (T2, T3) are identified with T1's practice because it is performed by a peer and (e) projection mechanism runs: assessors imaginarily project their own possible decisions and practice on assessee's performance. Thus, both refractions join in the middle and a shared space of reflection (f) on teaching experiences, possibilities and understanding is increased. Lastly, T2 and T3 use assessment criteria (g) to base their assessment on and generate peer feedback (h) about T1's performance. However, this feedback speaks not only about T1's performance, but also about T2 and T3's projection on it and their teaching experience, as in the quote:

> *I think we can easily drop to criticize a lesson, saying 'you know, this is good and this is wrong'. But when you start comparing yourself with, seeing your own lessons, you could say 'you know that I made the same mistakes but I did not realize'. And from this work you can say 'maybe I should have used another concept, or I should have done something else.' (F2, T13:35)*

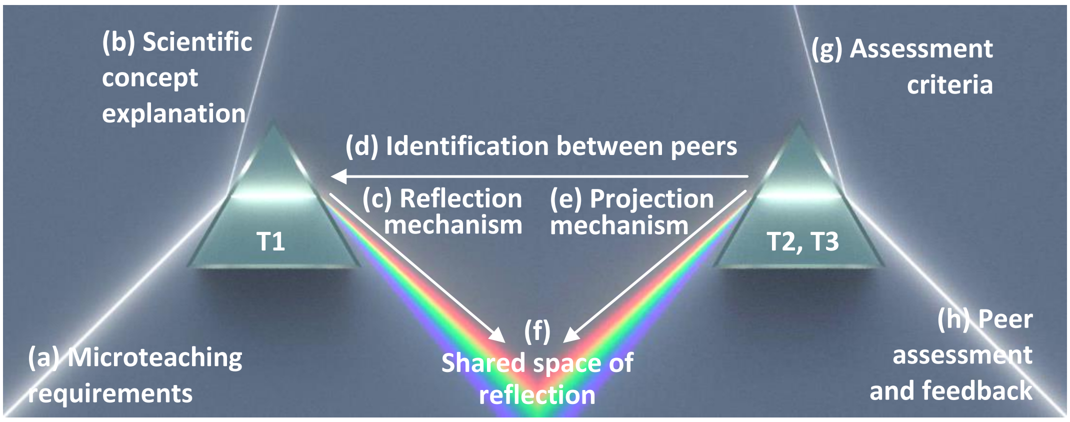


Figure 20: Model of projection and reflection mechanisms in Peer Assessment

In addition, considering the synthesis that Topping (2005) presented as a theoretical model of peer assisted learning, this prior conceptualization might imply a contribution to one of his points. He indicated that peer assisted learning can extend or restructure the declarative knowledge, modify current capabilities and rebuild new understanding, which leads to a joint construction of shared understanding between the helper and the helped, situated in their context of application and adapted to the idiosyncrasies in their perceptions. The current research embodied this part of the conceptualization. PA of the quality of explanations was based on the teachers' joint construction of shared meaning about the elements of quality. This was done exploring and transforming their implicit theories into assessment criteria. Likewise, peer feedback and the incorporation of peer's suggestions allowed teachers having the sense of successful practice, leading to consolidation as described in Topping's model. Thus, it was expected the generalization from the specific situated practice into wider alternatives he mentioned, which happens in this case when teachers started their first eventual job teaching in real schools. From the researcher's perspective, in teacher education between the co-construction of meaning and the concrete improvements, the two mechanisms mentioned would be situated, opening the spectrum of teachers' possibilities, experiences and understanding. This idea would enrich the model application of PA in ITE.

Finally, the rubric this research used contributed to understand "quality of teacher conceptual explanations in science" construct from literature review criteria, which is modelled on Figure 21. In this model, clarity is achieved when the description of scientific concept is given before its definition, when each part of the explanation conducts to the next one and has cohesive ties. All parts are coherent with the concept and show relations with the others, and they are related with pupils' experiences. In high quality explanations the parts are presented in a progressive sequence and the main ideas included in the explanation scaffold the concept construction. A high quality explanation it is accurate and sufficient in terms of scientific understanding but it might contain generalizations or simplifications that are beneficial for pupils' learning. Finally, four support clusters can be used in a good quality explanation: (1) analogies, metaphors, models or simulations, (2) examples, images, graphs, experiments or demonstrations, (3) gestures or voice inflections to stress and represent the concept and (4) illustration of common mistakes in the understanding of the concept or using pupils' errors as an opportunity for learning.

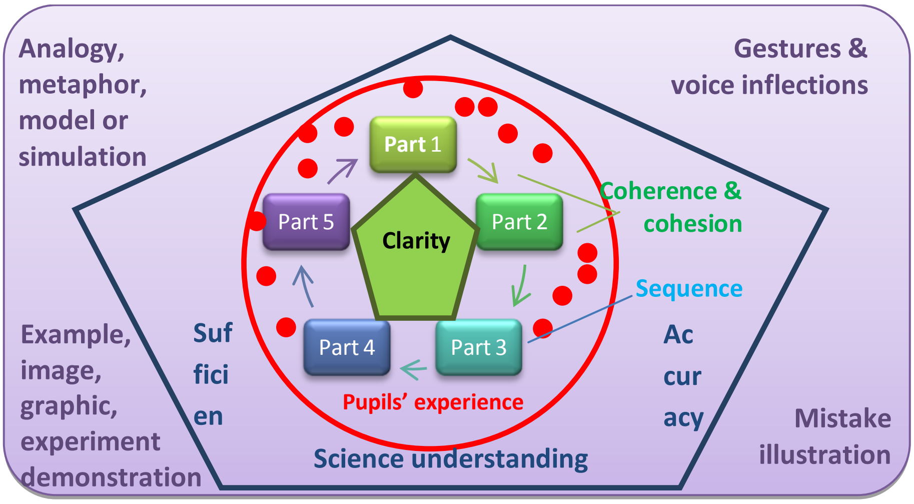


Figure 21: Model of quality criteria of science teacher explanations

(Based on literature review).

6. Conclusions and action implications

This chapter discusses the main conclusions regarding the research aims, highlighting the contribution of the current study to research through a summary of the findings according to the research questions. Then, some unexpected results uncovered by the initial questions are presented. Finally, implications of the findings for new research, practice and policy are offered in order to project this study into other research fields or possibilities.

6.1. Conclusions

The present research explored the development of the skill of explaining scientific concepts in Chilean pre-service teachers during initial teacher education (ITE). Its importance was identified in view of the fact that explanation is the strategy science Chilean teachers use more frequently during their lessons (Preiss et al., 2012). Also, according to Alvarado (2012) Chilean pupils considered knowing how to explain was the most important characteristic of a good teacher, and the Ministry of Education has identified and stressed teachers' weakness in this area (Gobierno de Chile, 2010a, 2013). Otherwise, a few studies reported using peer assessment (PA) in initial science teacher education (Gess-Newsome & Lederman, 1990; Hume, 2012; Kilic & Cakan, 2007; M. Kim, 2009; Sluijsmans, Brand-Gruwel, & van Merriënboer, 2002; Tsai et al., 2002; Wen & Tsai, 2008).

Thus, an answer to this question was sought: to what extent does PA facilitate change in Chilean pre-service science teachers' conceptions and practice to explain scientific concepts? This research identified pre-service teachers' strengths and weaknesses in the practice to explain scientific concepts in the experimental groups and proved it is possible to develop the skill in the context of PA during ITE. A change in student teachers' conceptions was illustrated in this research, and it has also identified which elements of the skill of explaining were easier and more difficult to change and transfer into real teaching contexts. The factors associated with this change, the facilitators and the obstacles for the transference of skill were described from the student teachers' and researcher's perspectives. A more detailed view of the answers to the specific research questions is presented in the following pages.

(1) Are the conceptions about the quality of explanations of pre-service science teachers exposed to PA different from those of teachers who have not been exposed to PA?

Before the PA, experimental and comparison groups showed in the PA questionnaire a similar distribution of evaluative comments into the categories General Aspects (GA), Knowledge Aspects (KA), Pedagogical Knowledge Aspects (PKA), Pedagogical Content Knowledge Aspects (PCKA) which were used to analyse their conceptions. Here, the category most used to assess the quality of explanations was PKA. This distribution remained relatively stable in the control group measured after the intervention period. However, the experimental group showed a change in this proportion: most of the comments in the experimental group were related to PCKA, while in the control group the majority of comments continued being related to PKA. This implied reorganization in experimental group teachers' conceptions. This finding followed the same pattern when the feedback sessions were analysed. Before the intervention, student teachers' justifications were mostly related to PKA, then GA, KA and finally PCKA. However, at the end of the intervention most of the student teachers' comments in the experimental group were related to PCKA, followed by PKA, GA and then KA. Then, PKA and PCKA were inverted. Furthermore, after participating in PA 25% of the experimental group had a more critical judgement of the explanation used as a base-line measurement, while in control group it remained stable. This suggests PA contributed to gain criticism about the practice to explain.

(2) Are the implicit theories about the quality of explanations of pre-service teachers different according to their science knowledge?

To answer this question, the student teachers were asked to construct their own instruments as a device to assess their peers' performance. Through product analysis of the documents created, it was possible to observe how student teachers' theories varied according to the university they belonged to, then, probably the variations were explained by the different science knowledge they had. Although all the groups defined themselves as believing in a constructivist way of teaching science, at the moment of deciding why an explanation was better or not, the quantity of elements related to constructivist approach were very different. In this sense, University 1 and University 2 presented more elements than University 3, and in this last university the implicit theories about quality of explanations were less sophisticated, simpler and less articulated than in the others. This is coincident with their high, medium and low science knowledge, respectively.

(3) What are the explanation elements that pre-service science teachers use to explain scientific concepts? Are these elements equally modifiable when using PA?

To answer this question, the construction of an instrument to evaluate science teachers' explanations was needed. After development, the rubric was found to be a valid instrument to assess different elements that were part of the skill of explaining scientific concepts. These elements were: clarity, coherence and cohesion, sequence, accuracy, sufficiency, connection with pupils' experience, using metaphors, analogies, simulations or models, using examples, demonstrations, experiments, images or graphs, using gestures and voice inflections and illustrating pupils' errors or common mistakes as a source of learning. The instrument allowed identifying pre-service teachers' strengths (sequence and using examples demonstrations, experiments, images or graphs) and weaknesses (using metaphors, analogies, simulations or models and errors or common mistakes for learning) in their skill. The instrument was sensitive enough to detect changes in student teachers' performance after participating in the PA intervention and also the elements that were the most resistant to change. After PA, the experimental group improved their explanations rated against the rubric, and the difference between their performance pre and post intervention was statistically significant ($p<.05$). Teachers who started at a lower level improved most at the end of PA. In general, clarity of explanation and connecting the explanation with pupils' experience were the easiest elements to modify, while using metaphors, analogies, simulations or models and using errors or common mistakes as a source or learning were the most difficult ones.

(4) Is it possible to change conceptions and practice about explaining concepts using PA?

Results from PA questionnaires and feedback sessions suggested that participants' conceptions changed towards the development of PCK during the intervention in. As detailed above, their practice was also possible to improve significantly and largely.

(5) What are the main elements associated with the change process?

The factors involved in the subjective process of change were described from the point of view of the student teachers and the researcher. Student teachers considered the reflection on their own practice as the engine of their conceptions and practice modification. This reflection was nurtured by challenging their previous theories with their current practice through microteaching episodes, which allowed them to visualize their strengths and weaknesses and analyse them from a changed focus of analysis. Their focus of analysis was changed through a negotiation of meaning process when they constructed the assessment criteria to evaluate their peers' teaching practice. This process enhanced pre-service teachers' self-regulation and improvement of their practice. Student teachers also mentioned systematic evaluation of their practice, knowing new criteria and the presence of an empathetic facilitator as intervenient conditions that affected positively the changes. As mentioned before, from the point of view of the researcher, there were two mechanisms involved in the subjective process of change that have their nature in social constructivism of meaning. These were reflection and projection, following the analogy of a mirror proposed by l'Anson et al. (2003). The peers' practice reflected what most of student teachers would do in a similar teaching situation and teachers in the role of assessor were identified and imaginarily projecting their own possible decision making process on their peers' practice.

(6) Do the changes in explanation practice sustain over the time after PA?

Most of the participants performed explanations with a better score in the follow-up study than in the pre and even post measurement. Only one participant maintained his post-test score which was already high. This implies the improvement obtained after PA was possible to be sustained six months after the intervention. Regarding the explanations elements, the two more difficult aspects to improve using PA were the same that were not possible to transfer from teacher education context (using metaphors, analogies, simulations or models and illustrating pupils' errors or common mistakes as a source of learning) into real teaching practice, whereas the other eight elements were maintained or potentiated.

(7) What are the factors (facilitators and obstacles) affecting the transference of good practice to explain scientific concepts into real teaching?

Regarding the transference process from simulated teaching in university into real teaching contexts, the follow-up interviews conducted with the teachers showed facilitators, obstacles and mediators from the participant teachers' point of view. They linked the transference process with the confidence they currently had at the time of teaching. Firstly, having a good tutor or guide teacher was the most salient facilitator. Then, the participants mentioned counting with support from the school and flexibility to implement lesson according to their objectives. Also, PA intervention was seen as a positive factor for the transference process especially because of the criteria construction and the formative critique. Otherwise, among the obstacles, participant teachers stressed the lack of time for planning and reflection they had at school, the low interest in science education in the school, pupils' lack of science knowledge and lack of participation in lessons, and their own weak and disconnected initial teacher education received. The mediator factors were understood receiving influences from facilitators and obstacles, and then affecting positively or negatively the transference of skills process and teachers' confidence. The participant teachers identified their own knowledge of science and pupils, their classroom climate and school resources provision as mediators. The three groups of factors interact, giving a different subjective experience to beginning teachers that can enhance or/and make difficult at the same time their skill transference. Nevertheless, participant teachers in the follow-up study maintained or potentiated the advance they had at the end of PA intervention, which might imply that although they recognised more obstacles than facilitators for the generalisation, they were able to cope with those and construct good quality explanations.

Besides, there were serendipitous discoveries in this research. One of the main purposes was to explore pre-service teachers' conceptions about the quality of explanations and their possible change using PA. It was expected a modification from simple conceptions to deeper and richer ones as a result of PA. However, student teachers' conceptions about explaining scientific concepts changed during PA not only in deepness but also in the applicability of their pedagogical knowledge into content knowledge. This finding led to assume a transformation of their conceptions towards the construction of PCK. The student teachers' initial conceptions were mostly about pedagogical knowledge, but at the end these were applied to specific science knowledge, showing this as an unexpected result. Likewise, the quality of teachers' explanations according to the rubric was not directly associated with their prior science knowledge because the three groups of student teachers were equally heterogeneous at the beginning of PA, despite the very different science education they received during the teaching career. Although it was not a sough relation, this was interesting to report as a serendipitous discovery. This is not coincident with the study of Sevian and Gonsalves (2008) where it was stated that the extent of an explainers' knowledge of the science research topic was strongly influenced the quality of the explanation.

6.2. Action implications

6.2.1. Implications for further research

In terms of immediate further research, this study might broaden the area of PA in ITE, for example introducing in PA interventions in teaching with more than one re-teaching process after the assessment. Another research has found that three rounds of PA can improve the quality of the products that science pre-service teachers created (Tsai et al., 2002). Although the current research found this result can be achieved in two rounds, it would be interesting to look at the possible effects of three or more PA rounds in teacher skills development but also in their sustainability. Would an increase of PA rounds lead to a major advance in the skills or in time they would be sustained? Special attention should be paid to the maximum of PA rounds if using microteaching, trying to avoid overloading the teachers and determining an efficient combination of number of sessions and achievement. For example, Sparks and McCallon (1974) found that six sessions of re-teaching in microteaching were associated with more negative attitudes towards microteaching due to teachers' feeling of overwork.

Furthermore, it would be interesting to look at different effects of PA or peer feedback combined with other techniques derived from the critical analysis of student teachers' practice. For instance, comparing groups receiving peer feedback only, with others groups having peer and/or tutor feedback and also self-assessment can be relevant to understand how different sources of formative critique might conduct to different results modifying pre-service teachers' practice. In this argument, it is recommended to have an experimental research design including random assignation of teachers to experimental and control groups, to be able to establish causality between the variables and PA or other type of assessment, and possible interactions among these variables.

Otherwise, this research has established the characteristics of good teacher explanations according to criteria found in the literature review, but the link between those good explanations identified with students' learning outcomes is not yet established. It would be valuable to study children's learning outcomes after receiving good explanations, comparing this achievement with other learners who would receive less quality explanations. The focus of the instrument usage in this research was entirely formative, but it is possible to project a trend likely to measure teacher performance jointly with students' results to determine its effectiveness. Thus, an interesting question for further research would be which characteristics of teacher explanations identified within the rubric are more effective in terms of students' learning outcomes?

Finally, it is recommendable for future research in the context of Chilean ITE to explore the extent to which teacher education programmes are developing the key competences good teachers should have (i.e. creating PCK to teach or skills of explaining). This might lead to an investigation of the current standards of Chilean teacher education, establishing comparisons between universities according to their effectiveness in the implementation process and their outcomes. However, any intention would not be successful without a careful consideration of how higher education operates in Chile: receiving students with extremely heterogeneous knowledge and skills, and usually having a very limited budget, human resources and time for engaging in research. Addressing and overcoming those issues is recommended to avoid difficulties and creating more funding possibilities for research in ITE, as it has been proposed but as yet unsuccessful (García-Huidobro, 2011; Gobierno de Chile, 2005).

6.2.2. Implications for practice

Considering the broad scope of this research, several implications for practice have been identified beyond the context of Chilean ITE.

First of all, the rubric constructed by the researcher can be considered as a tool for improving diagnostic and intervention mechanisms in initial teacher competences, in this case, the skill to explain scientific concepts. This is relevant because there are just a few validated instruments to have an evaluation of teaching in action (Pianta, La Paro, & Hamre, 2007). Several different instruments are currently used in ITE institutions (for example, Teaching Performance Assessment Form from Ekiti State University used in Oluwatayo and Adebule (2012) or other university forms investigated in Koziol, Minnick, and Sherman (1996)) but little examination of teacher skills acquisition has been done during ITE (Koziol et al., 1996) . In this sense, this rubric can be useful to orientate decision making in curricular activities to develop student teachers' skills before they are faced with real teaching practice.

Likewise, as this rubric has proved to be sensitive enough to detect even small changes when they exist and also inform when students' skills are not being modified, it can be a device to design progress evaluations and contribute to creation of measurable goals in teacher education. Having in mind the conceptualization of Day et al. (2007) about the professional pathways in teacher career[2], it is necessary to have more and better information about the student teachers' progress in the first years of real teaching, in order to detect early stops, expected and unexpected advances. With the information this rubric gives it is possible to create peer learning groups monitoring the development of other student teachers' skills, facilitating student mentoring in one of the most important phases where the commitment with teaching and having support from colleagues are clues to face the challenges.

The possible mentoring between student teachers who are identified with this rubric in a better level of skill of explaining can have important repercussions in terms of cost-effectiveness, considering supervising teachers' practice is often a very costly part of teacher education programmes in terms of time, human and material resources (Roth & Tobin, 2001). The current research can serve as an example to design innovative methodologies similar to peer formative assessment that consider the educational and constructive peers' power, which also distributes responsibility of teaching-learning among the learners. This can help student teachers to assume a more professional role in their teaching and own learning.

As a final point, modification in teachers' conceptions based on their implicit theories as presented in this research is a good input to project sustainable interventions in time. It is known that maintaining good practice is complex when this practice is not supported by deep dispositions that underpin or give it sustenance. In this argument, the results obtained in this research can be considered deep and durable because they were oriented to explore and transform student teachers' representations about quality into richer and more situated knowledge. Thus, it is recommended for other programmes that intend to change specific

[2] Six professional life phases were identified by these researchers based on teachers' experiences at school: 0-3 years -commitment: support and challenge, 4-7 years – Identity and efficacy in the classroom, 8-15 years – Managing changes in role and identity, 16-23 years – Work-life tensions: Challenges to motivation, 24-30 years Challenges to sustaining motivation and 31+ years Sustaining/declining motivation, coping with change, looking to retire.

student teachers' behaviours, patterns or practice to start first de-constructing the theories that support practice in order to re-construct them. This was valued by the participants in this research, and it could be a source of sustainability for other projects.

> *I think our ability to create an instrument was very important because it helps us to improve our own practices. Then, from what we have created, we correct ourselves now. (I5:3)*

6.2.3. Implications for policies of teacher education

In terms of policy implications, it is important to note that currently in Chile a particular policy program with extended theoretical underpinnings is nearly to be approved by the Senate. This program aims to redesign the teacher career by establishing differentiated steps in teacher professional development. The three main changes proposed are: increasing the requirements of the universities to accept student teachers; implement a registration system at the end of the initial training; increasing the salary of beginning teachers according to their performance and consolidating the assessment of in-service teacher practices in order to orientate the salary distribution (García-Huidobro, 2011). Even though these reforms seem positive, this redefinition is not considering the role of ITE to assure good teaching practice construction or the transition process from university to the real schools yet. In this aspect, this research stressed two points that are useful to consider in the proposal: the need of valid instruments to assess teachers' skills development during ITE based on standards and the crucial role that tutor or guide teachers plays supporting beginning teachers in their transition process. Indeed, in this research, beginning teachers mentioned that although they can learn how to design and deliver good practice in teacher education (for instance explaining), they need an experienced teacher to support them implementing such practice in classroom.

In this sense, the work presented here could stimulate the policy debate in the accountability of teacher education institutions regarding the development of competences and the assurance that they are transferred into real teaching context. This study aimed to develop a competence impacting its main components: cognitive, behavioural and transference. Here the skill was tested before, after the intervention and also followed-up in its application to the real context, which is the complete cycle to assure the competence is acquired. It could be taken as a model to design accountability policies looking for institutions responsibility.

The scale of the debate about teacher education is extensive and multifaceted nowadays. To generate national policy strategies and develop pre-service teacher education, there is a need for more studies of innovations that aim to achieve pre-service teacher learning during their time at university. Exploring PA in ITE could help not only Chilean teacher education policies but also other countries around the world that have taken teacher education as a national commitment. In this way, it would be valuable to investigate which methodologies are being used by teacher education programmes across the world to diagnose teachers' practical strengths and weaknesses before going to the labour market. Here, it would be useful to offer the rubric designed in this research as a valid instrument to identify and intervene. Possible adaptations for other subject areas apart from science might be a powerful further development of the instrument looking for its generalization.

Also, some comparisons can be established between countries, for example, modifying the rubric created in this research according to the orientation of different countries' programmes or applying it in a first attempt without modifications and comparing the student teachers' achieved level. Likewise, it would be interesting to question how might in-service teacher quality of explanations be measured with this rubric? Would expert/novice teachers have a stable or variable pattern according to the content being taught or their countries? Exploring these areas would allow having a larger sample of teachers to apply more rigorous statistical analysis in testing the instrument, and also expanding its scope of utility. Moreover, it would help understanding how teachers' professionalism is developed in different cultures and how it might affect students' learning outcomes considering teaching as a situated practice. These understandings may lead to design better teacher induction and retention policies.

For instance, a particular policy program recently approved in Chile was the introduction of professional standards in teacher education programmes. It has been received as an icon of expected success in the standardization of the existent varied programmes and in the evaluation of their quality (García-Huidobro, 2011). However, evidence from several studies, (Cofré et al., 2010; Gobierno de Chile, 2005; Vergara & Cofré, 2008) have indicated that initial teacher education has a low impact in teachers' practice and conceptions despite the past reformation attempts. This means the institutions are not yet able to take their responsibility of modifying student teachers' entry representations and skills (Gobierno de Chile, 2005).

In this area, the results of this research can have important implications for redesigning teaching careers profiles based on competences as it is happening now not only in Chile but also in other countries such as Venezuela, Argentina, Brazil, Spain or Mexico that are in a similar educational level and facing the same challenges in teacher education. In this area, this research results can inform the evaluation of teachers' practice by their peers, teacher educators and also in national proficiency tests. Taking this point into account, the findings of this research are valuable also for developed countries that nowadays are using standards for teacher education programmes such as Australia, Scotland or the United States. These results might offer a valid way to assess how specific standards can progress along the time, which transcends the interest of specific countries to broader contexts.

7. References

ACARA. (2011). Australian Curriculum, Assessment and Reporting Authority. *The Australian Science Curriculum,* (18th November 2011). Retrieved from http://www.australiancurriculum.edu.au/Science/Rationale

Akerson, V. L., & Hanuscin, D. L. (2007). Teaching nature of science through inquiry: Results of a 3-year professional development program. *Journal of Research in Science Teaching, 44*(5), 653-680.

Alvarado. (2012). Resultados de la Encuesta "Mis Profesores y Yo". *Noticias PUCV.* Retrieved 29th August, 2012, from http://prensa.ucv.cl/?p=14441

Alvarado, Cabezas, G., Falck, D., & Ortega, M. E. (2012). *La Evaluación Docente y sus instrumentos: discriminación del desempeño docente y asociación con los resultados de los estudiantes*. Santiago de Chile: Ministerio de Educación.

Andreu, J. (2001). Las técnicas de análisis de contenido: una revisión actualizada. Retrieved from http://public.centrodeestudiosandaluces.es/pdfs/S200103.pdf

Appleton, K. (2006). Science Pedagogical Content Knowledge and Elementary School Teachers. In K. Appleton (Ed.), *Elementary Science Teacher Education* (pp. 31-54). New Jersey: Lawrence Erlbaum Associates Inc. Publishers.

Appleton, K. (2007). Elementary science teaching. In S. Abell & N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 493-535). Mahway: Lawrence Erlbaum.

Appleton, K., & Kindt, I. (2002). Beginning Elementary Teachers' Development as Teachers of Science. *Journal of Science Teacher Education, 13*(1), 43-61.

Arellano, E. L., Barcenal, T. L., Bilbao, P. P., Castellano, M. A., Nichols, S., & Tippins, D. J. (2001). Case-based pedagogy as a context for collaborative inquiry in the Philippines. *Journal of Research in Science Teaching, 38*(5), 502-528.

Arzi, H. J., & White, R. T. (2008). Change in teachers' knowledge of subject matter: A 17-year longitudinal study. *Science Education, 92*(2), 221-251.

Ashton, P. (1984). Teachers' Sense of Efficacy: A Self-or Norm-Referenced Construct? *Florida Journal of Educational Research, 26*(1), 29-41.

ASTA. (2002). *National Professional Standards for Highly Accomplished Teachers of Science*: Australian Science Teachers Association.

ASTA. (2009). *National Professional Standards for Highly Accomplished Teachers of Science*: Australian Science Teachers Association.

Avalos, B. (2003). La formación docente inicial en Chile. IESALC-UNESCO (Ed.) Retrieved from http://www.oei.es/docentes/info_pais/informe_formacion_docente_chile_iesalc.pdf

Avalos, B. (2011). Teacher professional development in Teaching and Teacher Education over ten years. *Teaching and Teacher Education, 27*(1), 10-20.

Bakkenes, I., Vermunt, J. D., & Wubbels, T. (2010). Teacher learning in the context of educational innovation: Learning activities and learning outcomes of experienced teachers. *Learning and Instruction, 20*(6), 533-548.

Bandura, A. (1995). Excercise of personal and collective efficacy in changing societies. In A. Bandura (Ed.), *Self-efficacy in changing societies* (pp. 1-45). New York: Cambridge University Press.

Bandura, A. (1997). *Self-efficacy: The exercise of control*: Worth Publishers.

Bardin, L. (1996). *Análisis de contenido*. Madrid: Akal ediciones.

Barnett, J., & Hodson, D. (2001). Pedagogical context knowledge: Toward a fuller understanding of what good science teachers know. *Science Education, 85*(4), 426-453.

Barrow, L. (2006). A Brief History of Inquiry: From Dewey to Standards. *Journal of Science Teacher Education, 17*(3), 265-278.

Beck, J., Czerniak, C. M., & Lumpe, A. T. (2000). An Exploratory Study of Teachers' Beliefs Regarding the Implementation of Constructivism in Their Classrooms. *Journal of Science Teacher Education, 11*(4), 323-343.

Beck, R. J., King, A., & Marshal, S. K. (2002). Effects of videocase construction on preservice teachers' observations of teaching. *The Journal of experimental education, 70*(4), 345-361.

Bellezza, F. (1996). Mnemonic methods to enhance storage and retrieva. In E. Bjork & R. Bjork (Eds.), *Memory* (2nd ed., pp. 345-380). Orlando, FL: Academic Press.

Bencze, & Bowen, G. M. (2009). A national science fair: Exhibiting support for the knowledge economy. *International Journal of Science Education, 31*(18), 2459-2483.

Bencze, Hewitt, J., & Pedretti, E. (2001). Multi-media case methods in pre-service science education: Enabling an apprenticeship for praxis. *Research in Science Education, 31*(2), 191-209.

Berry, A., & Loughran, J. (2010). *What do we know about effective CPD for developing science teachers' pedagogical content knowledge?* Paper presented at the International Seminar, Profesional Reflections, National Science Learning Center, York.

Bird, T., Anderson, L., Sullivan, B., & Swidler, S. A. (1992). *Pedagogical balancing acts: A teacher educator encounters problems in an attempt to influence prospective teachers' beliefs*. Washington DC: Office of Educational Research and Improvement.

Bitar, S. (2011). *Formación docente en Chile.* (1st ed.). Santiago de Chile: PREAL.

Boling, E. C. (2007). Linking technology, learning, and stories: Implications from research on hypermedia video-cases. [Article]. *Teaching and Teacher Education, 23*(2), 189-200.

Borman, K. M., Mueninghoff, E., Cotner, B. A., & Frederick, P. B. (2009). Teacher preparation programs. In L. J. Saha & A. G. Dworkin (Eds.), *International Handbook of Research on Teachers and Teaching* (Vol. 21, pp. 123-140): Springer US.

Bostock, S. (2000). Student peer assessment. *The Higher Education Academy* Retrieved 25th July 2012, from www.keele.ac.uk/depts/aa/landt/lt/docs/bostock_peer_assessment.htm

Boud, D. (1990). Assessment and promotion of academic values. *Studies in Higher Education, 15*, 110-113.

Boud, D. (1999). Situating academic development in professional work: using peer learning. *International Journal for Academic Development, 4*(1), 3-10.

Boud, D., Cohen, R., & Sampson, J. (1999). Peer Learning and Assessment. *Assessment & Evaluation in Higher Education, 24*(4), 413-426.

Boyatzis, R. (1998). *Transforming qualitative information: Thematic analysis and code development*. Thousands Oaks, CA: Sage.

Braun, V., & Clarke, V. (2006). Unsing thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101.

Braund, M. (2008). *Starting science... again? Making progress in Science Learning*. London: SAGE.

Braund, M., & Reiss, M. (Eds.). (2004). *Learning science outside the classroom*. London: Routledge.

Bravo, D., Falck, D., González, R., Manzi, J., & Peirano, C. (2009). Do Good Teachers Make a Difference? The link between teacher certification programs and student achievement in Chile. *LACEA*.

Brewer, W. F., Chinn, C. A., & Samarapungavan, A. (2000). Explanation in scientist and children. In F. C. Keil & R. A. Wilson (Eds.), *Explanation and Cognition* (pp. 279-298). Cambridge, MA: The MIT Press.

Brookhart, S. M., & Freeman, D. J. (1992). Characteristics of Entering Teacher Candidates. *Review of Educational Research, 62*(1), 37-60.

Brown, G. A., & Daines, J. M. (1981). Can explaining be learnt? Some lecturers' views. *Higher Education, 10*(5), 573-580.

Bryan, L. A., & Abell, S. K. (1999). Development of professional knowledge in learning to teach elementary science. *Journal of Research in Science Teaching, 36*(2), 121-139.

Buckley, B. C. (2000). Interactive multimedia and model-based learning in biology. *International Journal of Science Education, 22*(9), 895-935.

Butefish, W. L. (1990). Science Teachers' Perceptions of Their Interactive Decisions. *The Journal of Educational Research, 84*(2), 107-114.

Calderhead, J., & Robson, M. (1991). Images of teaching: Student teachers' early conceptions of classroom practice. *Teaching and Teacher Education, 7*(1), 1-8.

Campanario, J. M., & Moya, A. (1999). ¿Cómo enseñar ciencias? Principales tendencias y propuestas. *Enseñanza de las Ciencias, 17*(2), 179-192.

Canal de Leon, P. (2000). El análisis didáctico de la dinámica del aula: tareas, actividades y estrategias de enseñanza. In F. Palacios & P. Canal de Leon (Eds.), *Didáctica de las ciencias experimentales*. España: Marfil.

Carlsen, W. S. (1991). Effects of new biology teachers' subject-matter knowledge on curricular planning. *Science Education, 75*, 631-647.

Carlsen, W. S. (1993). Teacher knowledge and discourse control: Quantitative evidence from novice biology teachers' classrooms. *Journal of Research in Science Teaching, 30*(5), 471-481.

Carr, M., Baker, M., Bell, B., Jones, A., Kirkwood, V., Pearson, J., & Symington, D. (1994). The constructivist paradigm and some implications for science content and pedagogy. In P. J. Fenshamp, R. F. Gunstone & R. T. White (Eds.), *The Content of Science: a Constructivist Approach to its Teaching and Learning* (pp. 147-160). London: Falmer Press.

Carter, C. (1990). Teachers' knowledge and learning to teach. In W. R. Houston (Ed.), *Handbook of research on teacher education* (pp. 291-310). New York: Macmillan.

Carter, K., & Doyle, W. (1996). Persional narrative and life history in learning to teach. In J. Sikula, T. Buttery & E. Guyton (Eds.), *Handbook of research on teacher education* (2nd ed.). London: Prentice Hall International.

Catalán, J. (2010). *Teorías subjetivas: aspectos teóricos y prácticos*. La Serena: Editorial Universidad de La Serena.

Clark, C. M. (1988). Asking the Right Questions About Teacher Preparation: Contributions of Research on Teacher Thinking. *Educational Researcher, 17*(2), 5-12.

Clarke, D., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and Teacher Education, 18*, 947-967.

Cobern, W., & Loving, C. (2000). Defining "science" in a multicultural world. *Science Education, 85*, 50-67.

Cobern, W., & Loving, C. (2002). Investigation of preservice elementary teachers' thinking about science. *Journal of Research in Science Teaching, 39*(10), 1016-1031.

Cofré, H. (Ed.). (2010). *Cómo mejorar la enseñanza de las ciencias en Chile: Perspectivas internacionales y desafíos nacionales*. Santiago de Chile: Ediciones Universidad Católica Silva Henríquez.

Cofré, H., Camacho, J., Galaz, A., Jimenez, J., Santibanez, D., & Vergara, C. (2010). La educación científica en Chile: Debilidades de la enseñanza y futuros desafíos de la educación de profesores de ciencia. *Estudios Pedagógicos, 36*(2), 289-303.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale: Lawrence Erlbaum Associates.

Cole, V., John-Steiner, V., Scribner, S., & Souberman, U. E. (Eds.). (1978). *L.S. Vygotsky. Mind in society: The development of higher processes.* Cambridge, MA: Harvard University Press.

Conner, L. (2010). Using metamatrices to analyse learning from multiple data sources. In S. Rodrigues (Ed.), *Using Analytical Frameworks for Classroom Research* (Vol. 1). London: Routledge.

Cook, M. P. (2006). Visual representations in science education: The influence of prior knowledge and cognitive load theory on instructional design principles. *Science Education, 90*(6), 1073-1091.

Copeland, W. D. (1975). The relationship between microteaching and student teacher classroom performance. *Journal of educational research, 68*(8), 289-293.

Crawford, B. A. (1999). Is It Realistic to Expect a Preservice Teacher to Create an Inquiry-based Classroom? *Journal of Science Teacher Education, 10*(3), 175-194.

Czerniak, C., & Lumpe, A. (1996). Relationship between teacher beliefs and science education reform. *Journal of Science Teacher Education, 7*(4), 247-266.

Dagher, Z. (1992). Analysis and synthesis of studies related to the effectiveness of analogies in science learning. *Paper presented at the annual meetng of the National Association for Research in Science Teaching*, Boston, M.A.

Dagher, Z., & Cossman, G. (1992). Verbal explanations given by science teachers: Their nature and implications. *Journal of Research in Science Teaching, 29*(4), 361-374.

Daniels, H. (2004). Activity theory, discourse and Bernstein. *Educational Review, 56*(2), 121-132.

Danielson, C. (1996). *Enhancing Professional Practice. A Framework for Teaching*. Alexandria, Virginia: Association for Supervision and Curriculum Development.

Danielson, C. (2011). *The framework for teaching evaluation instrument* (1st ed.). Princeton: The Danielson group.

Danto, A. C. (1985). *Narration and knowledge*. New York: Columbia University Press.

Darling-Hammond, L., & Hammerness, K. (2002). Toward a pedagogy of cases in teacher education. *Teaching Education, 13*(2), 125-135.

Darling-Hammond, L., Wise, A. E., & Klein, S. P. (1995). *A License to Teach. Building a Profession for 21st-Century Schools*. San Francisco: Westview Press.

Davis, E. A., & Petish, D. (2005). Real-World Applications and Instructional Representations Among Prospective Elementary Science Teachers. *Journal of Science Teacher Education, 16*(4), 263-286.

Davis, E. A., Petish, D., & Smithey, J. (2006). Challenges new science teachers face. *Review of Educational Research, 76*(4), 607-651.

Dawes, L. (2004). Talk and learning in classroom science. *International Journal of Science Education, 26*, 677-695.

Dawson, P. (1975). Effect of video feedback on teacher behavior. *Journal of educational research, 68*(5), 197-201.

Day, C., Stobat, G., Sammons, P., Kingston, A., Gu, Q., & Smees, R. (2007). Variations on Teachers' Work, Lives and Effectiveness. London: Institute of Education.

De Jong, O., Van Driel, J. H., & Verloop, N. (2005). Preservice teachers' pedagogical content knowledge of using particle models in teaching chemistry. *Journal of Research in Science Teaching, 42*(8), 947-964.

Denley, P., & Bishop, K. (2007). *Learning Science Teaching: Developing a Professional Knowledge Base*: Open university press.

Denzin, N., & Lincoln, Y. (Eds.). (1994). *Handbook of qualitative research* Thousand Oaks: Sage.

Do-Yong, P. (2006). *Curriculum Reform Movement in the US – Science Education* Paper presented at the 1st Pacific Rim Conference on Education, Hokkaido, Japan. http://www.hokkyodai.ac.jp/international-c/conference/SS2-8_Do-Yong_Park.pdf

Docentemas. (2010). The national teacher evaluation system Retrieved 09th november, 2010, from www.docentemas.cl

Druit, R. (1999). Conceptual change approaches in science education. In W. Schnotz, S. Vosniadou & M. Carretero (Eds.), *New perspectives on conceptual change* (pp. 263-282). Amsterdam, NL: Pergamon.

Duit, R., & Treagust, D. F. (2003). Conceptual change: A powerful framework for improving science teaching and learning. *International Journal of Science Education, 25*(6), 671 - 688.

Duncan, R., Pilitsis, V., & Piegaro, M. (2010). Development of Preservice Teachers' Ability to Critique and Adapt Inquiry-based Instructional Materials. *Journal of Science Teacher Education, 21*(1), 1-14.

Dupin, J., & Joshua, S. (1989). Analogies and modeling analogies in teaching: Some examples in basic electricity. *Science Education, 73*(2), 207-224.

Duran, D., & Monereo, C. (2008). The Impact of Peer Tutoring on the Improvement of Linguistic Competence, Self-Concept as a Writer and Pedagogical Satisfaction. *School Psychology International, 29*, 481-499.

Dweck, C. S., Chiu, C.-y., & Hong, Y.-y. (1995). Implicit Theories: Elaboration and Extension of the Model. *Psychological Inquiry, 6*(4), 322-333.

Dwyer, C. A. (1994). *Development of a Knowledge Base for the PRAXIS III: Classsroom Performance Assessment Criteria*. Princeton. NJ: Educational Testing Service.

Dye, J., Schatz, I., Rosenberg, B., & Coleman, S. (2000). Constant Comparison Method: A Kaleidoscope of Data. *The qualitative report.* Retrieved 30th may, 2011, from http://www.nova.edu/ssss/QR/QR4-1/dye.html

Eder, M. L. (2005). La explicación en la enseñanza de las ciencias. *Enseñanza de las Ciencias, Numero extra* (VII Congreso Internacional).

Edgington, J. (1995). *How research physicists and high-school physics teachers deal with the scientific explanation of a physical phenomenon.* Paper presented at the Annual Meeting of the National Association for Research in Science Teaching. April 22-25, San Francisco, CA.

Edgington, J. (1997). *What constitutes a scientific explanation?* Paper presented at the Annual meeting of the National Association for Research in Science Teaching, Oak Brook, IL. http://eric.ed.gov/PDFS/ED406190.pdf

Eick, C. J., & Reed, C. J. (2002). What makes an inquiry-oriented science teacher? The influence of learning histories on student teacher role identity and practice. *Science Education, 86*(3), 401-416.

Enochs, L. G., & Riggs, I. M. (1990). Further development of an elementary science teaching efficacy belief instrument: A preservice elementary scale. *School Science and Mathematics, 90*(8), 694-706.

Eshach, H. (2006). *Science literacy in primary schools and pre-schools*. Dordrecht, The Netherlands: Springer.

Euridyce Network. (2011). Science Education in Europe: National Policies, Practices and Research. Retrieved from http://eacea.ec.europa.eu/education/eurydice

Fang, Z. H. (1996). A review of research on teacher beliefs and practices. *Educational Research, 38*(1), 47-65.

Faye, J. (2009). Explanation and Interpretation in the Sciences of Man, from http://philsci-archive.pitt.edu/5187/

Feiman-Nemser, S. (2008). Teacher learning: How do teachers learn to teach? In M. Cochran-Smith, S. Feiman-Nemser, D. McIntyre & K. Demers (Eds.), *Handbook of research on teacher education. Enduring questions in changing contexts* (pp. 697-705). New York: Routledge.

Feiman-Nemser, S., & Buchmann, M. (1989). Describing Teacher Education: A Framework and Illustrative Findings from a Longitudinal Study of Six Students. *The Elementary School Journal, 89*(3), 365-377.

Feinstein, N. (2010). Salvaging science literacy. *Science Education, 95*(1), 168-185.

Fenshamp, P. J., Gunstone, R. F., & White, R. T. (Eds.). (1994). *The Content of Science: a Constructivist Approach to its Teaching and Learning.* London: Falmer Press.

Ferguson, R. (2008). If Multicultural Science Education Standards' Existed, What Would They Look Like? *Journal of Science Teacher Education, 19*(6), 547-564.

Feynman, R. P. (1994). *Six easy pieces*. Reading, MA: Helix Books.

Field, A. (Ed.). (2005). *Discovering statistics using SPSS* (2nd ed.). London: Sage.

France, B. (2010). Narrative interrogation. In S. Rodrigues (Ed.), *Using Analytical Frameworks for Classroom Research* (Vol. 1). London: Routledge.

Frisch, A. L., Camerini, L., Diviani, N., & Schulz, P. J. (2011). Defining and measuring health literacy: how can we profit from other literacy domains? *Health Promotion International, 27*(1), 117-126.

Furman, M. (2008). *Ciencias Naturales en la escuela primaria: colocando las piedras fundamentales del pensamiento científico*. Paper presented at the IV Foro Latinoamericano de Educación: Aprender y enseñar ciencias. Desafíos, estrategias y oportunidades, Buenos Aires, Argentina. http://coleccion.educ.ar/coleccion/CD23/contenidos/biblioteca/pdf/furman.pdf

Gagliardi, P. (Ed.). (1990). *Symbols and artifacts* (1st ed.). New York: Aldine de Gruyter.

García-Huidobro, J. (2011). La política docente hoy y la formación de profesores. *Docencia, 43*, 12-42.

Geddis, A. N. (1993). Transforming subject-matter knowledge: the role of pedagogical content knowledge in learning to reflect on teaching. *International Journal of Science Education, 15*(6), 673 - 683.

Geelan, D. (2003). Teacher expertise and explanatory frameworks in a successful physics classroom. *Australian Science Teachers Journal, 49(3)*(3), 22-32.

Geelan, D. (2009). *Explaining topics in physics: An international video study*. Paper presented at the World Conference on Educational Multimedia, Hypermedia and Telecommunications 2009, Honolulu, HI, USA. http://www.editlib.org/p/31839

Geelan, D. (2012). Teacher explanations. In B. J. Fraser (Ed.), *Second international handbook of science education* (pp. 987-999). Dordrecht, The Netherlands: Springer.

Gess-Newsome, J. (1999). Pedagogical content knowledge: an introduction and orientation. In J. Gess-Newsome & N. G. Lederman (Eds.), *Examining pedagogical content knowledge* (pp. 3-17). Boston, M.A.: Kluwer.

Gess-Newsome, J., & Lederman, N. G. (1990). The preservice microteaching course and science teachers' instructional decisions: A qualitative analysis. *Journal of Research in Science Teaching, 27*(8), 717-726.

Gess-Newsome, J., & Lederman, N. G. (1993). Preservice biology teachers' knowledge structures as a function of professional teacher education: A year-long assessment. *Science Education, 77*(1), 25-45.

Gibson, S., & Dembo, M. H. (1984). Teacher efficacy: A construct validation. *Journal of educational psychology, 76*(4), 569.

Gielen, S., Dochy, F., & Onghena, P. (2010). An inventory of peer assessment diversity. *Assessment & Evaluation in Higher Education, 36*(2), 137-155.

Gil, D. (1994). Relaciones entre conocimiento escolar y conocimiento científico. *Investigación en la escuela, 23*, 17-32.

Ginns, I. S., & Watters, J. J. (1999). Beginning Elementary School Teachers and the Effective Teaching of Science. *Journal of Science Teacher Education, 10*(4), 287-313.

Glaser, B., & Strauss, A. L. (1967). *The discovery of grounded theory*. Chicago, IL: Aldine Publishing Company.

Glynn, S., Taasoobshirazi, G., & Fowler, S. (2007). Analogies: explanatory tools in web-based science instruction. *Educational technology magazine, 47*(5), 45-50.

Gobierno de Chile. (2001). *Estándares de desempeño para la formacion inicial de docentes*. Santiago de Chile: Ministerio de Educación. Retrieved from http://www.educarchile.cl/Userfiles/P0001/File/estandares_formacion_docentes.pdf.

Gobierno de Chile. (2003). *Marco para la buena enseñanza*. Santiago de Chile: Ministerio de Educación.

Gobierno de Chile. (2005). *Informe comisión sobre formación inicial docente*. Santiago de Chile: Ministerio de Educación.

Gobierno de Chile. (2007). *What does PISA tell us about the education of high school students in Chile? New analyses and perspectives on the PISA 2006 results.* Santiago de Chile: Ministerio de Educación.

Gobierno de Chile. (2009a). *Objetivos Fundamentales y Contenidos Mínimos Obligatorios de la Educación Básica y Media. Actualización 2009*. Santiago de Chile: Ministerio de Educación.

Gobierno de Chile. (2009b). *Orientaciones para la planificación de la enseñanza*. Santiago de Chile: División de Educación General. Ministerio de Educación.

Gobierno de Chile. (2010a). *Resultados evaluación docente 2009*. Santiago de Chile: Ministerio de Educación.

Gobierno de Chile. (2010b). *Resumen de resultados PISA 2009 Chile*. Santiago de Chile: Ministerio de Educación.

Gobierno de Chile. (2011a). *Acreditación de la educación superior*. Valparaíso: Retrieved from http://www.bcn.cl/guias/acreditacion-de-la-educacion-superior.

Gobierno de Chile. (2011b). *Estandáres orientadores para egresados de carreras de pedagogía en educación básica*. Santiago de Chile: Ministerio de Educación.

Gobierno de Chile. (2012a). *Bases curriculares 2012 ciencias naturales Educación Básica.* Santiago de Chile: Ministerio de Educación.

Gobierno de Chile. (2012b). *Estándares orientadores para carreras de educación parvularia*. Santiago de Chile: Ministerio de Educación.

Gobierno de Chile. (2012c). *Estandáres orientadores para carreras de pedagogía en educación media.* Santiago de Chile: Ministerio de Educación.

Gobierno de Chile. (2012d). *Informe nacional de resultados SIMCE 2011*. Santiago de Chile: Ministerio de Educación.

Gobierno de Chile. (2013). *Resultados nacionales de la evaluación docente 2012*. Santiago de Chile: Ministerio de Educación.

González, B., & Moreno, T. (1998). *Las analogías en la enseñanza de las Ciencias*. Paper presented at the Actas II simposio sobre la docencia de las ciencias experimentales en la enseñanza secundaria.

González, C., Martínez, M. T., Martínez, C., Cuevas, K., & Muñoz, L. (2009). La educación científica como apoyo a la movilidad social: Desafíos en torno al rol del profesor secundario en la implementación de la indagación científica como enfoque pedagógico. *Estudios pedagógicos (Valdivia), 35*, 63-78.

Goodwin, A. J. (1995). Understanding secondary school science: A perspective of the graduate scientist beggining teacher. *School Science Review, 76*(276), 100-109.

Graber, K. C. (1996). Influencing student beliefs: The design of a "high impact" teacher education program. *Teaching and Teacher Education, 12*(5), 451-466.

GTE. (2006). *The Standard for full registration*. Scotland: General Teaching Council.

Gunstone, R. F., Slattery, M., Baird, J. R., & Northfield, J. R. (1993). A case study exploration of development in preservice science teachers. *Science Education, 77*(1), 47-73.

Guskey, T. R. (1982). Differences in teachers' perceptions of personal control of positive versus negative student learning outcomes. *Contemporary Educational Psychology, 7*(1), 70-80.

Gutierrez, K. D., Baquedano-Lopez, P., Tejeda, C., & Rivera, A. (1999). Building a culture of collaboration through hybrid language practices. *Theory into practice, 38*, 87-93.

Hagger, H., Burn, K., Mutton, T., & Brindley, S. (2008). Practice makes perfect? Learning to learn as a teacher. *Oxford Review of Education, 34*(2), 159-178.

Halliday, M., & Hasan, R. (1976). *Cohesion in English*. London: Longman.

Harford, J., & MacRuairc, G. (2008). Engaging student teachers in meaningful reflective practice. *Teaching and Teacher Education, 24*(7), 1884-1892.

Hermans, R., van Braak, J., & Van Keer, H. (2008). Development of the Beliefs about Primary Education Scale: Distinguishing a developmental and transmissive dimension. *Teaching and Teacher Education, 24*(1), 127-139. doi: http://dx.doi.org/10.1016/j.tate.2006.11.007

Hoey, M. (1991). *Patterns of lexis*. Oxford: Oxford University Press.

Holt-Reynolds, D. (1992). Personal History-Based Beliefs as Relevant Prior Knowledge in Course Work. *American Educational Research Journal, 29*(2), 325-349.

Horwood, R. H. (1988). Explanation and description in science teaching. *Science Education, 72*(1), 41-49.

Hume, A. C. (2012). Primary Connections: Simulating the classroom in initial teacher education. *Research in Science Education 42*, 551-565.

IBM Corp. (2010). IBM SPSS statistics for Windows (Version 19.0.). Armonk, NY: IBM Corp.

Isikoglu, N., Basturk, R., & Karaca, F. (2009). Assessing in-service teachers' instructional beliefs about student-centered education: A Turkish perspective. *Teaching and Teacher Education, 25*(2), 350-356.

James, M. C., & Scharmann, L. C. (2007). Using analogies to improve the teaching performance of preservice teachers. *Journal of Research in Science Teaching, 44*(4), 565-585.

Jindal-Snape, D., & Topping, K. (2010). Observational analysis within case study design. In S. Rodrigues (Ed.), *Using Analytical Frameworks for Classroom Research* (Vol. 1). London: Routledge.

Justi, R. (2006). Enseñanza de las ciencias basada en la elaboración de modelos. *Enseñanza de las Ciencias, 24*(2), 173-184.

Kallenbach, W., & Gall, M. (1969). Microteaching versus conventional methods in training rlementary intern teachers. *Journal of Educational Research, 63*(3), 136-141.

Kane, P. R., & Temple, E. (1997). Who is a good teacher? Who is a good school head? *Independent School, 56*, 50-57.

Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science*. Cambridge: MIT Press.

Karmiloff-Smith, A., & Inhelder, B. (1974). If you want to get ahead, get a theory. *Cognition, 3*(3), 195-212.

Kilic, G. B., & Cakan, M. (2007). Peer assessment of elementary science teaching skills. *Journal of Science Teacher Education, 18*(1), 91-107.

Kim, H., & Hannafin, M. (2008). Grounded design of web-enhanced case-based activity. *Educational Technology Research and Development, 56*(2), 161-179.

Kim, J. (1995). Explanation. In R. Audi (Ed.), *The Cambridge dictionary of philosophy* (pp. 256-257). Cambridge, England: Cambridge University Press.

Kim, M. (2009). The impact of an elaborated assessee's role in peer assessment. *Assessment & Evaluation in Higher Education, 34*(1), 105-114. doi: 10.1080/02602930801955960

King, K. P., & Wiseman, D. L. (2001). Comparing Science Efficacy Beliefs of Elementary Education Majors in Integrated and Non-Integrated Teacher Education Coursework. *Journal of Science Teacher Education, 12*(2), 143-153.

Klassen, R. M., Bong, M., Usher, E. L., Chong, W. H., Huan, V. S., Wong, I. Y. F., & Georgiou, T. (2009). Exploring the validity of a teachers' self-efficacy scale in five countries. *Contemporary Educational Psychology, 34*(1), 67-76. doi: http://dx.doi.org/10.1016/j.cedpsych.2008.08.001

Kollar, I., & Fischer, F. (2010). Peer assessment as collaborative learning: A cognitive perspective. *Learning and Instruction, 20*(4), 344-348. doi: http://dx.doi.org/10.1016/j.learninstruc.2009.08.005

Koziol, S., Jr., Minnick, J. B., & Sherman, M. (1996). What student teaching evaluation instruments tell us about emphases in teacher education programs. *Journal of Personnel Evaluation in Education, 10*(1), 53-74. doi: 10.1007/bf00139469

Kozma, R. (2003). The material features of multiple representations and their cognitive and social affordances for science understanding. *Learning and Instruction, 13*(2), 205-226.

Kpanja, E. (2001). A study of the effects of video tape recording in microteaching training. *British Journal of Educational Technology, 32*(4), 483-486.

Krause, M. (2005). *Psicoterapia y cambio. Una mirada desde la subjetividad.* Santiago: Ediciones Universidad Catolica de Chile.

Krause, M., Cornejo, M., & Radovic, J. (1997). *Diseño de estudios cualitativos*. Santiago: Ministerio de Salud.

Krause, M., De la Parra, G., Arístegui, R., Dagnino, P., Tomicic, A., Valdés, N., . . . Ben-Dov, P. (2007). The evolution of therapeutic change studied through generic change indicators. *Psychotherapy Research, 17*(6), 673 - 689.

Kukanja, M. (2007). Peer observation and assessment of teaching. *Pedagoska Obzorja, 22*(1-2), 178-190.

Kurz, T. L., & Batarelo, I. (2010). Constructive features of video cases to be used in teacher education. *TechTrends, 54*(5), 46-53.

l'Anson, J., Rodrigues, S., & Wilson, G. (2003). Mirrors, Reflections and Refractions: The contributions of microteaching to reflective practice. *European Journal of Teacher Education, 26*(2), 189-200.

Latorre, M. (2003). En el ojo del huracán: ¿qué características tiene el saber pedagógico en uso de nuestros profesores?, ¿Es posible contribuir desde la formación inicial? *Boletín de Investigación Educacional, Pontificia Universidad Catolica de Chile, 18*, 285-296.

Lawson, M. J., Askell-Williams, H., & Murray-Harvey, R. (2009). Dimensions of quality in teacher knowledge. In L. J. Saha & A. G. Dworkin (Eds.), *International Handbook of Research on Teachers and Teaching* (Vol. 21, pp. 243-257): Springer US.

Lebak, K., & Tinsley, R. (2010). Can Inquiry and Reflection be Contagious? Science Teachers, Students, and Action Research. *Journal of Science Teacher Education*, 1-18.

LeCompte, M. D. (2000). Analyzing Qualitative Data. *Theory Into Practice, 39*(3), 146.

Lederman, N. G. (1999). Teachers' understanding of the nature of science and classroom practice: Factors that facilitate or impede the relationship. *Journal of Research in Science Teaching, 36*(8), 916-929.

Lederman, N. G., & Gess-Newsome, J. (1989). *A qualitative analysis of the effects of a microteaching course on preservice science teachers' instrucional decisions and beliefs about teaching*. Paper presented at the 62nd Annual meeting of the National Association for Research in Sciencie teaching, San Francisco, CA.

Lederman, N. G., & Zeidler, D. L. (1987). Science teachers' conceptions of the nature of science: Do they really influence teaching behavior? *Science Education, 71*(5), 721-734.

Leite, L., Mendoza, J., & Borsese, A. (2007). Teachers' and prospective teachers' explanations of liquid-state phenomena: A comparative study involving three European countries. *Journal of research in science teaching, 44*(2), 349-374.

Limon, M., & Carretero, M. (1997). Las ideas previas de los alumnos, ¿qué aporta este enfoque a la enseñanza de la ciencias? *Construir y enseñar las ciencias experimentales* (pp. 19-46). Argentina: Aique.

Lin, S. S. J., Liu, E. Z., Chiu, C. H., & Yuan, S.-M. (2001). Peer review: an effective web-learning strategy with the learner as both adapter and reviewer. *IEEE Transactions on Education, 44*(3), 246-251.

Lloyd, J. K., Smith, R. G., Fay, C. L., Khang, G. N., Wah, L. L. K., & Sai, C. L. (1998). Subject knowledge for science teaching at primary level: a comparison of pre-service teachers in England and Singapore. *International Journal of Science Education, 20*(5), 521-532.

Lokan, J., Hollingsworth, H., & Hackling, M. (2006). *Teaching science in Australia. Results from the TIMSS 1999 video study.* Melbourne: ACER.

Lortie, D. (1975). *Schoolteacher: A Sociological Study*. Chicago: University of Chicago Press.

Lortie, D. (2002). *Schoolteacher: A Sociological Study* (2nd ed.). Chicago: University of Chicago Press.

Loughran, J. (1994). Bridging the gap: An analysis of the needs of second-year science teachers. *Science Education, 78*(4), 365-386.

Loughran, J., Berry, A., & Mulhall, P. (2007). Pedagogical Content Knowledge: What Does it Mean to Science Teachers? Contributions from Science Education Research. In R. Pintó & D. Couso (Eds.), (pp. 93-105): Springer Netherlands.

Loughran, J., & Russell, T. (Eds.). (2002). *Improving Teacher Education Practices Through Self-study*. London: Routledge-Falmer.

MacArthur, C. A., Schwartz, S. S., & Graham, S. (1991). Effects of a reciprocal peer revision strategy in special education classrooms. *Learning Disabilities Research and Practice, 6*(201-210).

Maclean, R., & White, S. (2007). Video reflection and the formation of teacher identity in a team of pre-service and experienced teachers. *Reflective Practice: International and Multidisciplinary Perspectives, 8*(1), 47 - 60.

Magnusson, S., Krajcik, J., & Borko, H. (2002). Nature, Sources, and Development of Pedagogical Content Knowledge for Science Teaching. In J. Gess-Newsome & N. G. Lederman (Eds.), *Examining Pedagogical Content Knowledge* (Vol. 6, pp. 95-132). Netherlands: Springer.

Mansour, N. (2009). Science teachers' beliefs and practices: issues, implications and research agenda. *International Journal of Environmental & Science Education, 4*(1), 25-48.

Martin, R., Sexton, C., & Gerlovich, J. (2002). *Teaching science for all children: methods for constructing understanding*. Boston: Allyn and Bacon.

Marzano, R., Pickering, D., & Pollack, J. (2001). *Classroom interactions that works*. Alexandria, VA: ASCD Press.

Mayer, R. E., & Gallini, J. K. (1990). When is an illustration worth ten thousand words? *Journal of educational psychology, 82*, 715-726.

Mayer, R. E., & Jackson, J. (2005). The case for coherence in scientific explanations: Quantitative details can hurt qualitative understanding. *Journal of applied experimental psychology, 11*(1), 13-18.

Mayo, A., Sharma, M., & Muller, D. (2009). Qualitative Differences Between Learning Environments Using Videos in Small Groups and Whole Class Discussions: A Preliminary Study in Physics. *Research in Science Education, 39*(4), 477-493.

McDiarmid, G. W. (1990). Challenging Prospective Teachers' Beliefs During Early Field Experience: A Quixotic Undertaking? *Journal of Teacher Education, 41*(3), 12-20.

McLaury, R. L. (2011). *Preservice science teacher beliefs about teaching and the science methods courses: Exploring perceptions of microteaching outcomes.* Ph.D. Thesis, Purdue University. Retrieved from http://udini.proquest.com/view/preservice-science-teacher-beliefs-pqid:2514714971/ ProQuest Dissertations And Theses database.

McLuckie, J., & Topping, K. J. (2004). Transferable skills for online peer learning. *Assessment & Evaluation in Higher Education, 29*(5), 563 - 584.

Meirink, J., Meijer, P., Verloop, N., & Bergen, T. (2009). Understanding Teacher Learning in Secondary Education: The Relations of Teacher Activities to Changed Beliefs about Teaching and Learning. *Teaching and Teacher Education, 25*(1), 89-100.

Mellado, V. (1998). The classroom practice of preservice teachers and their conceptions of teaching and learning science. *Science Education, 82*(2), 197-214.

Mellado, V. (2003). Cambio didáctico del profesorado de ciencias experimentales y filosofía de la ciencia. *Enseñanza de las Ciencias, 21*(3), 343-358.

Mestre, J. P., Dufresne, R. J., Gerace, W. J., Hardiman, P. T., & Touger, J. S. (1993). Promoting skilled problem-solving behavior among beginning physics students. *Journal of research in science teaching, 30*(3), 303-317.

Microsoft, C. (2010). Microsoft Office PowerPoint 2010 (Version 14). Seattle, WA: Microsoft.

Moghaddam, F., Harre, R., & Lee, N. (2008). Positioning and conflict: an introduction. In R. Moghaddam, R. Harre & N. Lee (Eds.), *Global conflict resolution trough positioning analysis* (pp. 3-21). New York: Springer.

Mohan, R. (2007). *Innovative science teaching for physical science teachers* (3rd ed.). India: Prentice Hall.

Moje, E., Collazo, T., Carrillo, R., & Marx, R. W. (2001). "Maestro, What is 'quality'?": Language, literacy and discourse in project-based science. *Journal of research in science teaching, 38*(4), 469-498.

Moreno, R., & Ortegano-Layne, L. (2008). Do classroom exemplars promote the application of principles in teacher education? A comparison of videos, animations, and narratives. *Educational Technology Research and Development, 56*(4), 449-465.

Muller, D. A., Sharma, M. D., & Reimann, P. (2008). Raising cognitive load with linear multimedia to promote conceptual change. [Article]. *Science Education, 92*(2), 278-296.

Munby, H., & Russel, T. (1992). Transforming chemistry research into chemistry teaching: the complexities of adopting new frames of experiences. In T. Russel & H. Munby (Eds.), *Teachers and teaching: from classroom to reflection* (pp. 87-104). New York: Falmer Press.

Murphy, P. K., Delli, L. A. M., & Edwards, M. N. (2004). The good teacher and good teaching: Comparing beliefs of second-grade students, preservice teachers, and inservice teachers. *The Journal of Experimental Education, 72*(2), 69-92.

National Research Council. (1996). *National Science Education Standards*. Washington DC: The National Academy Press.

Navarro, M., & Förster, C. (2012). Nivel de alfabetización científica y actitudes hacia la ciencia en estudiantes de secundaria: comparaciones por sexo y nivel socioeconómico. *Pensamiento Educativo, 49*(1), 1-17.

Neuendorf, A. (2001). *The content analysis guidebook*. Thousand Oaks, CA: Sage.

Newhouse, C. P., Lane, J., & Brown, C. (2007). Reflecting on Teaching Practices Using Digital Video Representation in Teacher Education. *Journal of Teacher Education, 3*(3).

Nicol, D. J., & Boyle, J. T. (2003). Peer instruction versus class-wide discussion in large classes: A comparison of two interaction methods in the wired classroom. *Studies in Higher Education, 28*(4), 457-473.

Norris, S. P., Guilbert, S. M., Smith, M. L., Hakimelahi, S., & Phillips, L. M. (2005). A theoretical framework for narrative explanation in science. *Science Education, 89*(4), 535-563.

NSTA. (2012). *Preservice Science Standards*. United States of America: National Science Teachers Association.

OECD. (2005). *Teachers Matter: Attracting, Developing and Retaining Effective Teachers*. Paris: OECD Publishing.

OECD. (2006). *Science Competences for Tomorrow's World. Vol 1 - Analysis*. Paris: OECD Publishing.

OECD. (2007). *PISA 2006: science competencies for tomorrow's world. Volume 1: Analysis.* Paris: OECD Publishing.

OECD. (2009). *PISA 2009 Assessment Framework. Key competencies in reading, mathematics and science*. Paris: OECD Publishing.

OECD. (2010a). PISA 2009 Results: Overcoming Social Background – Equity in Learning Opportunities and Outcomes (Volume II). Retrieved from http://www.oecd.org/pisa/pisaproducts/48852584.pdf

OECD. (2010b). PISA 2009 Results: What Students Know and Can Do – Student Performance in Reading, Mathematics and Science (Volume I) Retrieved from http://dx.doi.org/10.1787/9789264091450-en

Ogborn, J., Kress, G., Martins, I., & MGillicuddy, K. (1996). *Explaining Science in the Classroom*. London: Open University Press.

Ogborn, J., & Martins, I. (1996). Metaphorical understandings and scientific ideas. *International Journal of Science Education, 18*(6), 631 - 652.

Oluwatayo, J. A., & Adebule, S. O. (2012). Assessment of teaching performance of student-teachers on teaching practice. *International Education Studies, 5*(5), 109-115.

Onslow, B., Beynon, C., & Geddis, A. (1992). Developing a teaching style: A dilemma for student teachers. *Alberta Journal of Educational Research, 38*(4), 301-315.

Orlova, N. (2009). Video recording as stimulus for reflection in pre-service EFL teacher training. *English teacher forum, 2*, 30-35.

Orsmond, P., Merry, S., & Reiling, K. (1996). The importance of marking criteria in the use of peer assessment. *Assessment & Evaluation in Higher Education, 21*(3), 239-250.

Orsmond, P., Merry, S., & Reiling, K. (2000). The use of student derived marking criteria in peer and self-assessment. *Assessment & Evaluation in Higher Education, 25*(1), 23-38.

Osborn, J., & Dillon, J. (2008). *Science education in Europe: Critical reflections*. London: Nuffield Foundation.

Osborne, R., & Freyberg, P. (1985). *Learning in science: The implications of children's science*. Auckland: Heinemann.

Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research, 62*(3), 307-332.

Palma, C. (2012, 25th March 2012). Escasez de profesores de ciencias en Chile llega a 35%, *La Tercera*. Retrieved from http://diario.latercera.com/2012/03/25/01/contenido/pais/31-104633-9-escasez-de-profesores-de-ciencias-en-chile-llega-a-35.shtml

Patton, M. (2001). *Qualitative Evaluation and Research Methods* (3rd ed.). Newbury Park, CA: Sage Publications.

Pauline, R. F. (1993). Microteaching: An integral part of a science methods class. *Journal of Science Teacher Education, 4*(1), 9-17.

Phillips, D., & Hayes, B. (2008). Securing the oral tradition: reflective positioning and professional conversations in midwifery education. *Collegian, 15*, 109-114.

Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2007). *Classroom assessment scoring system (Observation Protocol Manual)*. Baltimore: Paul H. Brookes.

Podolefsky, N. S. (2007). Analogical scaffolding and the learning of abstract ideas in physics: Empirical studies. *Physical review special topics. Physics education research, 3*(2), 020104.

Porta, L., & Silva, M. (2003). La investigación cualitativa: El Análisis de Contenido en la investigación educativa. Retrieved from http://www.investigacioncualitativa.es/Paginas/Articulos/investigacioncualitativa/PortaSilva.pdf

Pozo, J., & Gomez, M. (2005). The Embodied Nature of Implicit Theories: The Consistency of Ideas about the Nature of Matter. *Cognition and instruction, 23*(3), 351-387.

Pozo, J., & Gómez, M. (1998). *Aprender a enseñar ciencias: Del conocimiento cotidiano al conocimiento científico*. Madrid: Ediciones Morata.

Pozo, J., Gomez, M., & Sanz, A. (1999). When change does not mean replacement: different representations for different contexts. In W. Schnotz, S. Vosniadou & M. Carretero (Eds.), *New Perspectives on Conceptual Change*. Oxford: Elsevier.

Preiss, D., Alegría, I., Espinoza, A. M., Núñez, M., & Ponce, L. (2012). *¿Cómo se enseña la ciencia en la escuela? Evidencia de un estudio audiovisual en aulas de escuelas públicas chilenas.* Paper presented at the Segundo congreso interdiscipinario de investigación en educación, Santiago, Chile.

QAA. (2000). *The Standard for Initial Teacher Education*. Scotland: Quality Assurance Agency for Higher Education.

QSR. (2011). NVivo qualitative data analysis software (Version 9): QSR International Pty Ltd.

Redman, C., & Fawns, R. (2010). How to use pronoun grammar analysis as a methodological tool for understanding the dynamic lived space of people. In S. Rodrigues (Ed.), *Using Analytical Frameworks for Classroom Research* (Vol. 1). London: Routledge.

Rice, D., & Roychoudhury, A. (2003). Preparing More Confident Preservice Elementary Science Teachers: One Elementary Science Methods Teacher's Self-Study. *Journal of Science Teacher Education, 14*(2), 97-126.

Richardson, V. (1996). The role of attitudes and beliefs in learning to teach. In J. Sikula (Ed.), *Handbook of research on teaching* (pp. 376-391). New York: Macmillan.

Rodrigo, M. J. (1985). Las teorías implícitas en el conocimiento social. *Infancia y Aprendizaje, 31*(32), 145-156.

Rodrigues, S. (2010). Exploring talk. Identifying register, coherence and cohesion. In S. Rodrigues (Ed.), *Using Analytical Frameworks for Classroom Research* (Vol. 1). London: Routledge.

Rodriguez, A. J. (2001). From gap-gazing to promising cases: Moving toward equity in urban education reform. *Journal of Research in Science Teaching, 38*, 1115-1129.

Rodríguez, A. J. (1993). El conflicto intergrupal desde las teorías implícitas. In M. J. Rodrigo, A. Rodríguez & J. Marrero (Eds.), *Las teorías implícitas. Una aproximación al conocimiento cotidiano.* (pp. 309-338). Madrid: Visor Distribuciones S.A.

Rodríguez, S., Núñez, J. C., Valle, A., Blas, R., & Rosario, P. (2009). Auto-eficacia Docente, Motivación del Profesor y Estrategias de Enseñanza. *Escritos de Psicología (Internet), 3*, 1-7.

Roth, W.-M. (in print). Making Use of Gestures, the Leading Edge in Literacy Development. In W. Saul (Ed.), *Communicating science: Examining the discourse*.

Roth, W.-M., & Tobin, K. (2001). Learning to teach science as practice. *Teaching and Teacher Education, 17*(6), 741-762.

Roth, W.-M., & Welzel, M. (2001). From activity to gestures and scientific language. *Journal of research in science teaching, 38*(1), 103-136.

Russel, T., & Munby, H. (1991). Reframing: The role of experiece in developing teachers' professional knowledge. In D. A. Schon (Ed.), *The reflective turn: Case studies in and on educational practice* (pp. 24-39). New York: Teachers' College Press.

Rutherford, F., & Ahlgren, A. (1990). Science for All Americans. *American Association for the Advancement of Science (AAAS)*. New York: Oxford University Press.

Ruys, I., Van Keer, H., & Aelterman, A. (2010). Collaborative learning in pre-service teacher education: an exploratory study on related conceptions, self-efficacy and implementation. *Educational Studies, 36*(5), 537-553. doi: 10.1080/03055691003729021

Ryan, G., & Bernard, H. (2000). Data management and analysis methods. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook fo qualitative research* (2nd ed., pp. 769-802). Thousand Oaks, CA: Sage.

Sampson, V., & Clark, D. (2007). Incorporating scientific argumentation into inquiry-based activities with online personally seeded discussions. *Science Scope, 30*(6), 43-47.

Sandoval, C. (2002). Módulo cuatro Investigación cualitativa. *Especialización en teoría, métodos y técnicas de investigación social*. Bogota: ICFES.

Sanmartí, N. (1994). Aprender ciencias desde la diversidad, ¿fuente de problemas o fuente de riqueza? *Aula de Innovación Educativa, 27*, 1-6. Retrieved from Aula Grao website: http://aula.grao.com/revistas/aula/027-actitudes-y-ciencias-naturales--asesoramiento-en-centros

Sanmartí, N. (2000). Enseñar y aprender ciencias: algunas reflexiones. In N. Sanmartí & R. M. Puyol (Eds.), *Guía práctica de ciencias de la naturaleza* (pp. 9-42). Barcelona: Praxis.

Sarantakos, S. (2005). *Social Research*. Hampshire: Palgrave Macmillan.

Scharmann, L., & Orth Hampton, C. (1995). Cooperative learning and preservice elementary teacher science self-efficacy. *Journal of Science Teacher Education, 6*(3), 125-133.

Schön, D. A. (1991). *The reflective practitioner: how professionals think in action*. Avebury: Ashgate Arena.

Schubert, W. H. (1992). Personal theorizing about teacher personal theorizing. In E. W. Ross, J. W. Cornett & G. McCutcheon (Eds.), *Teacher Personal Theorizing Connecting Curriculum Practice, Theory, and Research* (pp. 257-272). Albanym N.Y.: State University of New York Press.

Schwartz, D. L., Bransford, J. D., & Sears, D. (2005). Efficiency and innovation in transfer. In J. P. Mestre (Ed.), *Transfer of learning from a modern multidisciplinary perspective* (pp. 1-52). Charlotte, NC: Information Age Publishing.

Sevian, H., & Gonsalves, L. (2008). Analysing how scientists explain their research: A rubric for measuring the effectiveness of scientific explanations. *International Journal of Science Education, 30*(11), 1441-1467.

Sherin, M. G. (2005). Using video to support teachers' ability to notice classroom interactions. *Journal of technology and teacher education, 13*(3), 475-491.

Shin, N., Stevens, S., & Krajcik, J. (2010). Tracking student learning over time using construct-centered design. In S. Rodrigues (Ed.), *Using Analytical Frameworks for Classroom Research* (Vol. 1). London: Routledge.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*, 1-14.

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Educational Researcher, 15*(2), 4-14.

Shulman, L. S. (2006). More than competition. *Chronicle of Education Higher, 53*(2), 104.

Sillman, K., & Dana, T. M. (2001). Methaphor: A Tool for Promoting Prospective Elementary Teachers' Participation in Science Teacher Learning Community. *Journal of Science Teacher Education, 12*(2), 87-106.

Simmons, P. E., Emory, A., Carter, T., Coker, T., Finnegan, B., Crockett, D., . . . Labuda, K. (1999). Beginning Teachers: Beliefs and Classroom Actions. *Journal of research in science teaching, 36*(8), 930-954.

Sluijsmans, D., Brand-Gruwel, S., & van Merriënboer, J. J. G. (2002). Peer assessment training in teacher education: effects on performance and perceptions. *Assessment & Evaluation in Higher Education, 27*(5), 443-454.

Sluijsmans, D., Brand-Gruwel, S., van Merriënboer, J. J. G., & Bastiaens, T. J. (2002). The training of peer assessment skills to promote the development of reflection skills in teacher education. *Studies in Educational Evaluation, 29*(1), 23-42.

Sluijsmans, D., Brand-Gruwel, S., van Merriënboer, J. J. G., & Martens, R. L. (2004). Training teachers in peer-assessment skills: Effects on performance and perceptions. *Innovations in Education and Teaching International, 41*(1), 59-78.

Sluijsmans, D., & Prins, F. (2006). A conceptual framework for integrating peer assessment in teacher education. *Studies in Educational Evaluation, 32*(1), 6-22.

Smith, D. C. (2000). Content and Pedagogical Content Knowledge for Elementary Science Teacher Educators: Knowing our Students. *Journal of Science Teacher Education, 11*(1), 27-46.

Smith, E. (2008). Raising standards in American schools? Problems with improving teacher quality. *Teaching and Teacher Education, 24*(3), 610-622.

Snyder, J. L. (2000). An investigation of the knowledge structures of experts, intermediates and novices in physics. *International Journal of Science Education, 22*, 979-992.

Sonmez, D., & Can, M. H. (2010). *Preservice science teachers' ability to identify good teaching practices.* Paper presented at the Procedia - Social and Behavioral Sciences, Istanbul.

Southerland, S. A., & Gess-Newsome, J. (1999). Preservice teachers' views of inclusive science teaching as shaped by images of teaching, learning, and knowledge. *Science Education, 83*(2), 131-150.

Sparks, R. L., & McCallon, E. L. (1974). Microteaching: Its effect on student attitudes in an elementary science methods course. *Science Education, 58*(4), 483–487.

Spelman, B. J., & John-Brooks, C. (1972). Microteaching and Teacher Education: A Critical Reappraisal. *The Irish Journal of Education / Iris Eireannach an Oideachais, 6*(2), 73-92.

Star, J. R., & Strickland, S. K. (2008). Learning to observe: Using video to improve preservice teachers' ability to notice. *Journal of Mathematics teacher education 11*, 107-125.

Stofflett, R. T. (1994). The accommodation of science pedagogical knowledge: The application of conceptual change constructs to teacher education. *Journal of research in science teaching, 31*(8), 787-810.

Strauss, A. L. (1987). *Qualitative analysis for social scientists*. Cambridge: Cambridge University Press.

Strauss, A. L., & Corbin, J. M. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Newbury Park, CA: Sage Publications, Inc.

Strickland, J. F., Page, J. A., & Page, F. M. (1986). *Presevice teachers' perceptions of teaching as a career*. Paper presented at the Annual meeting of the Mid-South American Educational Research Association, Memphis, Tenesse.

Surgue, C. (1996). Student teachers' lay theories: Implications for professional development. In F. Goodson & A. Hargreaves (Eds.), *Teachers' professional lives*. London: Falmer.

Sweeney, A. E., Bula, O. A., & Cornett, J. W. (2001). The role of personal practice theories in the professional development of a beginning high school chemistry teacher. *Journal of Research in Science Teaching, 38*(4), 408-441.

Tabachnick, B. R., & Zeichner, K. M. (1999). Idea and action: Action research and the development of conceptual change teaching of science. *Science Education, 83*(3), 309-322.

Tamir, P., & Zohar, A. (1991). Anthropomorphism and teleology in reasoning about biological phenomena. *Science Education, 75*(1), 57-67.

Tann, S. (1993). Eliciting student teachers' personal theories. In J. Calderhead & P. Gates (Eds.), *Conceptualizing reflection in teacher development* (pp. 53-69). London: Falmer Press.

Taylor, S., & Bogdan, R. (1993). *Introduction to qualitative research methods: A guidebook and resource* (3rd ed.). New York: John Wiley & Sons, Inc.

Thompson, F., & Logue, S. (2006). An exploration of common student misconceptions in science. *International Education Journal, 7*(4), 553-559.

Tigchelaar, A., Brouwer, N., & Vermunt, J. D. (2010). Tailor-made: Towards a pedagogy for educating second-career teachers. *Educational Research Review, 5*(2), 164-183.

Topping, K. J. (2005). Trends in Peer Learning. *Educational Psychology, 25*(6), 631-645.

Topping, K. J. (2010). Peers as a source of formative assessment. In G. J. Cizek (Ed.), *Handbook of formative assessment* (pp. 61-74). New York: Routledge.

Topping, K. J., & Ehly, S. (1998). *Peer-assisted learning*. Mahwah, N.J.; London: Lawrence Erlbaum.

Treagust, D., & Harrison, A. (1999). The genesis of effective scientific explanations for the classroom. In J. Loughran (Ed.), *Researching teaching: Methodologies and practices for understanding pedagogy*. London: Routledge.

Trout, J. D. (2002). Scientific explanation and the sense of understanding. *Philosophy of Science, 69*, 212-233.

Trumbull, D. J., & Kerr, P. (1993). University researchers' inchoate critiques-of science teaching: Implications for the content of preservice science teacher education. *Science Education, 77*(3), 301-317.

Trumper, R. (2003). The need for change in elementary school teacher training--a cross-college age study of future teachers' conceptions of basic astronomy concepts. *Teaching and Teacher Education, 19*(3), 309-323.

Tsai, C., Lin, S. S. J., & Yuan, S.-M. (2002). Developing science activities through a networked peer assessment system. *Computers & Education, 38*(1–3), 241-252.

Tsangaridou, N. (2002). Enacted Pedagogical Content Knowledge in Physical Education: a Case Study of a Prospective Classroom Teacher. *European Physical Education Review, 8*(1), 21-36.

Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education, 17*(7), 783-805.

Tymms, P., Bolden, D., & Merrell, C. (2008). Science in English primary schools: trends in attainment, attitudes and approaches. *Perspectives on Education 1 (Primary Science)* (pp. 19-43).

Tytler, R. (2007). *Re-imagining Science Education Engaging students in science for Australia's future*. Australia: Australian Council for Educational Research Press.

Tytler, R., Osborne, J. F., Williams, G., Tytler, K., & Cripps, J. (2008). *Opening up pathways: Engagement in STEM across the primary-secondary school transition. A review of the literature concerning supports and barriers to Science, Technology, Engineering and Mathematics engagement at primary-secondary transition*. Canberra: Australian Department of Education.

UNESCO-OREALC. (2005). Habilidades para la vida a través de la educación científica. Santiago de Chile: XVII Reunion de Coordinadores Nacionales del LLECE.

UNESCO-OREALC. (2010). *Factores asociados al logro cognitivo de los estudiantes de América Latina y el Caribe*. Santiago, Oficina Regional de Educación para América

Latina y el Caribe de la UNESCO: Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación.

University of Dundee. (2007). Research Ethics: Code of Practice, from http://www.dundee.ac.uk/eswce/research/ethics.htm

Vaillant, D. (2009). Policies for Insertion into the Teaching Profession in Latin America: the pending debt. *Profesorado, 13*(1), 1-15.

Valanides, N. (2010). Analysis of interview data using the constant comparative analysis method. In S. Rodrigues (Ed.), *Using Analytical Frameworks for Classroom Research* (Vol. 1). London: Routledge.

Valencia, E., & Taut, S. (2011). Evidencia de inequidad en el acceso a los docentes mejor calificados en la Enseñanza Básica chilena Santiago de Chile: Centro de Medicion MIDEUC.

van Driel, J. H., & Abell, S. K. (2010). Science Teacher Education. In P. Peterson, E. Baker & B. McGaw (Eds.), *International Encyclopedia of Education* (3rd ed., pp. 712-718): Elsevier Science.

van Driel, J. H., Jong, O. D., & Verloop, N. (2002). The development of preservice chemistry teachers' pedagogical content knowledge. *Science Education, 86*(4), 572-590.

van Driel, J. H., Verloop, N., & De Vos, W. (1998). Developing science teachers' pedagogical content knowledge. *Journal of research in science teaching, 35*(6), 673-695.

van Peer, W., & Chatman, S. (2001). Introduction. In W. Van Peer & S. Chatman (Eds.), *New Perspectives on Narrative Perspective* (pp. 1-17). Albany, NY: State University of New York Press.

van Zundert, M., Sluijsmans, D., & van Merriënboer, J. J. G. (2010). Effective peer assessment processes: Research findings and future directions. *Learning and Instruction, 20*(4), 270-279. doi: http://dx.doi.org/10.1016/j.learninstruc.2009.08.004

Vergara, C., & Cofré, H. (2008). La enseñanza de las ciencias en la educación básica Chilena: Un camino por recorrer. *Revista Foro Educacional, 14*, 85-104.

Vermunt, J., & Endedijk, M. (2011). Patterns in teacher learning in different phases of the professional career. *Learning and Individual Differences, 21*, 294-302.

Vygotsky, L. S. (1986). *Thought and language*. Cambridge, MA: MIT Press.

Wallace, M. J. (1998). *Action research for language teachers*. Cambridge: Cambridge University Press.

Ward, H. J., Roden, C., Hewlett, C., & Foreman, J. (2008). *Teaching science in the primary clasroom*. London: SAGE.

Webb, N. M., Nemer, K. M., & Zuniga, S. (2002). Short circuits or superconductors? Effects of group composition on high-achieving students' science assessment performance. *American Educational Research Journal, 39*(4), 943-989.

Weinstein, C. (1989). Teacher education students' perceptions of teaching. *Journal of Teacher Education, 40*(2), 53-60.

Welch, W. (Ed.). (1979). *Twenty-five years of science curriculum development* (Vol. 7). Washington, DC: American Educational Research Association.

Wellington, J., & Ireson, G. (2008). *Science learning, science teaching*. London: Routledge.

Wen, M. L., & Tsai, C. (2008). Online peer assessment in an inservice science and mathematics teacher education course. *Teaching in Higher Education, 13*(1), 55-67.

Wenham, M. (2005). *Understanding primary science: ideas, concepts and explanations*. London: SAGE.

White, R. T. (1994). Dimensions of content. In P. J. Fenshamp, R. F. Gunstone & R. T. White (Eds.), *The Content of Science: A constructivist Approach to its Teaching and Learning*. London: Falmer Press.

Whitehead, J., & Fitzgerald, B. (2007). Experiencing and evidencing learning through self-study: New ways of working with mentors and trainees in a training school partnership. *Teaching and Teacher Education, 23*(1), 1-12.

Wideen, M., Mayer-Smith, J., & Moon, B. (1998). A Critical Analysis of the Research on Learning to Teach: Making the Case for an Ecological Perspective on Inquiry. *Review of Educational Research, 68*(2), 130-178.

Windschitl, M., Thompson, J., & Braaten, M. (2008). Beyond the scientific method: Model-based inquiry as a new paradigm of preference for school science investigations. *Science Education, 92*(5), 941-967.

Wong, D. E. (1993). Understanding the generative capacity of analogies as a tool for understanding. *Journal of research in science teaching, 30*, 1259-1272.

Woolfolk-Hoy, A., & Murphy, P. K. (2001). Teaching educational psychology to the implicit mind. In B. Torff & R. Sternberg (Eds.), *Understanding and teaching the intuitive mind* (pp. 145-186). Mahwah, NJ: Erlbaum.

Woolhouse, M. (1999). Peer Assessment: the participants' perception of two activities on a further education teacher education course. *Journal of Further and Higher Education, 23*(2), 211-219.

Wragg, E., & Brown, G. A. (2001). *Explaining in the secondary school*. London: Routledge.

Wu, H.-K., & Shah, P. (2004). Exploring visuospatial thinking in chemistry learning. *Science Education, 88*(3), 465-492.

Yerrick, R., Doster, E., Nugent, J. S., Parke, H. M., & Crawley, F. E. (2003). Social interaction and the use of analogy: An analysis of preservice teachers' talk during physics inquiry lessons. *Journal of Research in Science Teaching, 40*(5), 443-463.

Yerrick, R., Ross, D., & Molebash, P. (2005). Too Close for Comfort: Real-Time Science Teaching Reflections via Digital Video Editing. *Journal of Science Teacher Education, 16*(4), 351-375.

Zacharia, Z. C. (2005). The impact of interactive computer simulations on the nature and quality of postgraduate science teachers' explanations in physics. *International Journal of Science Education, 27*(14), 1741-1767.

Zanting, A., Verloop, N., & Vermunt, J. D. (2003). Using Interviews and Concept Maps to Access Mentor Teachers' Practical Knowledge. *Higher Education, 46*(2), 195-214.

Zeichner, K. M., & Gore, J. M. (1990). Teacher socialization. In W. R. Houston (Ed.), *Handbook of research on teacher education* (pp. 291-310). New York: Macmillan.

Zembal-Saul, C., Blumenfeld, P., & Krajcik, J. (2000). Influence of Guided Cycles of Planning, Teaching, and Reflection on Prospective Elementary Teachers' Science Content Representations. *Journal of Research in Science Teaching, 37*(4), 318-339.

Zembal-Saul, C., Krajcik, J., & Blumenfeld, P. (2002). Elementary student teachers' science content representations. *Journal of Research in Science Teaching, 39*(6), 443-463.

8. Appendixes

8.1. Introduction letter to Universities to negotiate participation

Letter sent to career leaders:

Dear XX . Dean of Education. University XX.

My name is Valeria Cabello, I am a Psychologist and Master in Educational Psychology from Pontificia Universidad Católica de Chile. I have worked several years in the instrument construction team of Docentemas project in the centre MIDE UC (as you can see in my CV in attachment). Nowadays I am in the United Kingdom doing doctoral studies in Educational Psychology at the University of Dundee. My research topic is the development of the skill to explain scientific concepts in science teachers during the initial teacher education based on peer assessment.

To carry out this research I need to develop a project with science student teachers in their last year of training. The project would enhance their links between theory and practice through peer assessment. This methodology has been proved to have very good results in other educational fields. The aim is to facilitate changes in teacher thinking that underpin an improvement in their teaching practice.

This is a challenge project which intends to re think the teacher education methodologies and create evidence-based knowledge which will be useful for your institution and students.

I would like to discuss with you the possibility to work with your science student teachers and the project design. I am attaching to this letter the research proposal which was already approved by my supervisors in University of Dundee. If you have any comment or suggestion I will be happy to integrate them into the final design.

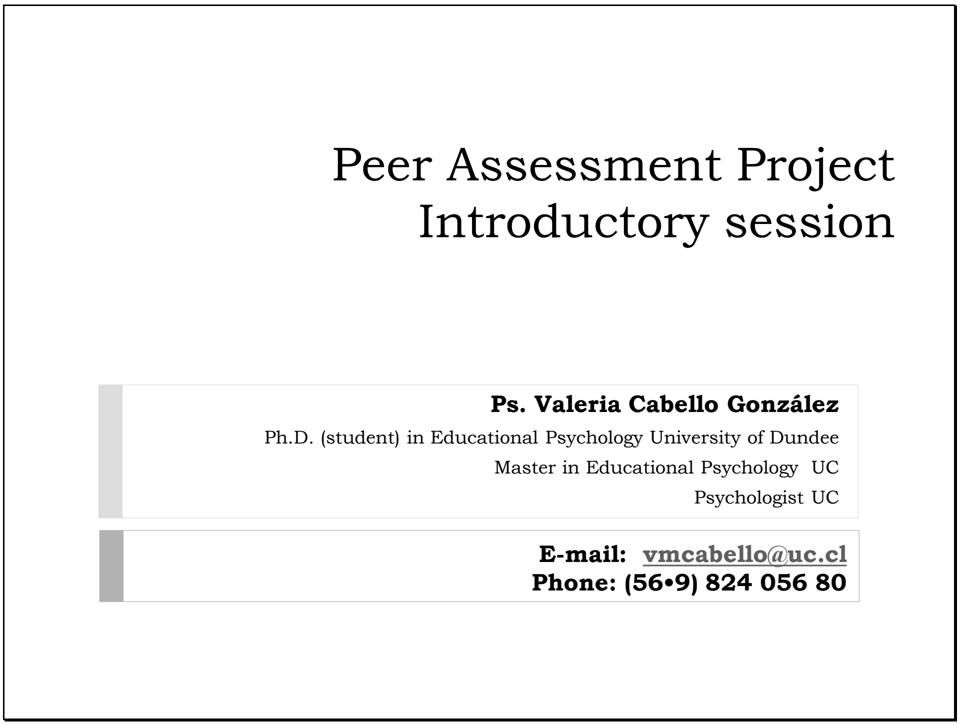
I hope this is the beginning of a fruitful work within your School of Education.

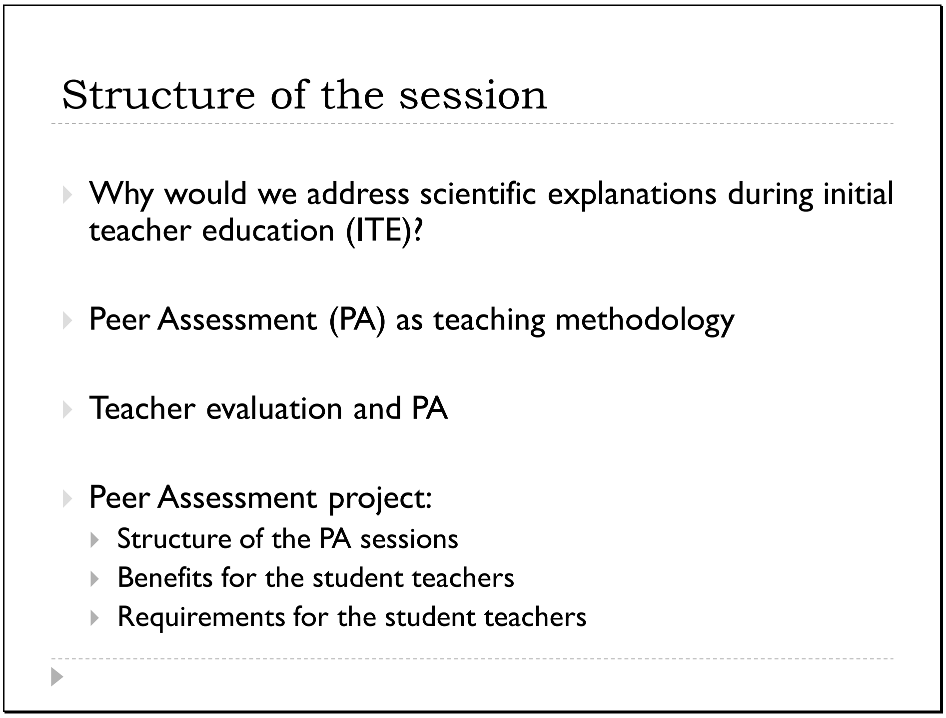
Sincerely,

--

Valeria M. Cabello González
Ph.D. student in Educational Psychology University of Dundee
Magíster en Psicología Educacional PUC
Psicóloga Pontificia Universidad Católica de Chile (PUC)
••
Fono: (44•7) 402 345 743
(56•9) 824 056 80
E-mail: v.m.cabellogonzalez@dundee.ac.uk
vmcabello@uc.cl
••
University of Dundee • School of Education
Scotland • United Kingdom

8.2. Presentation discussed with Deans of Education in negotiation meeting





Addressing scientific explanations during initial teacher education (ITE)

- Conceptual explanations are a teaching device used by the teacher to construct scientific concepts.
- In science there are several concepts that are difficult to understand, because of their high abstraction and students' concrete way of thinking.
- Concepts such as density, sedimentation, solubility, etc. are some exapmples of important scientific concepts children must construct to understand science.
- The teachers can define them verbally, demonstrate them, to present analogies or metaphors, to make gestures to represent them, among other tools to develop those scientific concepts.
- However, when do these tools are good?

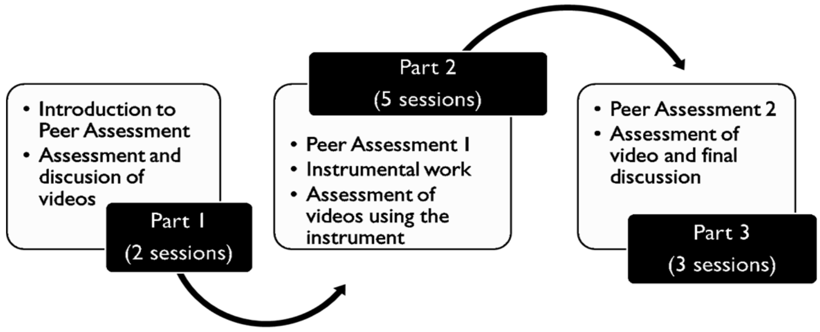
Peer Assessment (PA) as teaching methodology

- PA is an strategy generally used in evaluation of teaching quality by assessment systems in Chile.
- It has been also used as a teaching strategy in higher education such as in medical or science students.
- Basically a student or group presents a product which is formative evaluated by another student or group, giving feedback.
- Here it would be based on the formative evaluation principles and the teachers' critical reflection.
- Usually the roles change between evaluated and evaluator to complete the assessment for all.

Teacher evaluation and PA

- PA is used nowadays in the National teacher performance evaluation, to assess in a more reliable manner. It is considered as more fare than an evaluation from an external person.
- Evidence has shown that evaluated teachers feel more confident being evaluated by a peer.
- Likewise, the evaluator teachers value the process because for them it implies taking the other's perspective and learning new teaching techniques.
- It has been identified that evaluator teachers after playing this role improve their performance and face better the evaluation process.
- Research question: Would PA have similar positive effects when transferred into ITE?

Peer Assessment Project: Structure of the PA sessions



- Usually Peer Assessment 1 and 2 are taking 2 sessions each one, 30 mins. per teacher.

Peer Assessment Project: Benefits for the student teachers

- Knowing an internationally recognised methodology.
- Participation and pedagogycal innovation that might be transferred to future teaching.
- Better preparation to face different types of teacher performance evaluation.
- More knowledge about feedback techniques, and methodology evaluations.
- Developing the skill to assess others' practice which is highly required in the labour market.

Peer Assessment Project: Requirements for the student teachers

- To sign informed consent form giving authorisation to use their data for research purposes.
- To actively participate in all the sessions (including recorded microteaching and feedback sessions).
- Punctuality and respect for the peers' work. Motivation to learn how to assess and be assessed.
- Contact for possible project continuation.

8.3. Participant information sheet

INVITATION TO TAKE PART IN A RESEARCH STUDY

You are being asked to take part in a study, which is exploring a new teaching methodology in science teacher education. The researcher is Valeria Cabello, Ph.D. student in the School of Education, Social Work and Community Education (ESWCE) at University of Dundee, Scotland, United Kingdom. Professor Keith Topping and Professor Norman Reid are supervising the study.

PURPOSE OF THE RESEARCH STUDY

Your participation in this research would benefit you because you will learn a new teaching/assessing methodology that you can use in your future work as a teacher. Also, it will benefit your School of Education because they will introduce innovations in the curricula reported by international practice and international research. Moreover, this research may impact on the Chilean education system because its findings will be used to inform educational research and policy development.

TIME COMMITMENT

This study will use Peer Assessment (PA) as its main methodology. You will participate in a ten-session programme within and as part of your elective courses in your University. Therefore, it does not involve additional time demand. The project is going to start on August-September '11 and it will finish in November-December '11. First of all, you will receive brief training in PA using video tapes to illustrate it. In this period you will be asked to complete PA questionnaire and to evaluate the session process. Then, in a small group of classmates you will mutually assess your explanations in a microteaching episode (5-10 min) in order to create assessment criteria. These sessions will be observed by Valeria Cabello and you will be asked to give and receive feedback. Finally, you will use the assessment criteria to mutually assess your explanations in a second microteaching episode with your classmates and you will be asked to complete the same forms as in the first episode. After that, you will be called to participate in a group discussion to share your experiences within this programme. All the microteaching episodes will be video-audio recorded if you are in agreement. These digital records will facilitate data analysis of the verbal explanations and non-verbal communication you used when you were teaching during the observed sessions. If you do not agree to the video-audio recording, then you can decide to participate in the sessions but not in the data analysis. The final group discussion will be audio recorded. If you do not agree to the audio recording you can take part of the study but do not in the final discussion.

TERMINATION OF PARTICIPATION

You may decide to stop being a part of the research study at any time without explanation, and you will not incur any penalty.

RISKS

There are no known risks for you in this study.

COST, REIMBURSEMENT AND COMPENSATION

Your participation in this study is voluntary.

CONFIDENTIALITY/ANONYMITY

The data collected will not contain any personal information about you except a study number that will be given to each participant in order to match the different instruments and marks in previous science courses. All data records will be held for two years and saved on a password protected computer network, and a backup will be kept in a secure office in University of Dundee, ESWCE to avoid accidental loss or damage. After that time all data will be destroyed. All data generated will not be used for any other purpose than to inform this study and to disseminate information about this research. The only people with access to the data will be the researcher, supervisors and two assistants. If the results are published your identity will not be mentioned or identifiable, because pseudonyms will be used. However, please note, that under certain circumstances, under the law a researcher may have to reveal data that could reveal identities.

FOR FURTHER INFORMATION ABOUT THIS RESEARCH STUDY

Valeria Cabello Gonzalez will be glad to answer your questions about this study at any time personally or by phone calling to 08-2405680. If you want to find out about the final results of this study, you should ask for it to your Director of Education Career, or directly to the researcher Valeria Cabello Gonzalez: E-mail to v.m.cabellogonzalez@dundee.ac.uk or Mail: ESWE, University of Dundee, DD1 4HN, Dundee, Scotland, United Kingdom.

The University Research Ethics Committee of the University of Dundee has reviewed and approved this research study. The study will be conducted according to the "Research Ethics: Code of Practice". Please ask the researcher if you want a copy of this document.

8.4. Informed consent

By signing below you are agreeing that you have read and understood the Participant Information Sheet and that you agree to take part in this research study.

I agree to the video-audio recording of the programme sessions.
Please delete as appropriate YES NO

I agree to the audio recording of the final discussion.
Please delete as appropriate YES NO

________________________ ____________________________ ________________

Printed name of participant Participant's signature Date

Valeria Cabello González (researcher)
____
Printed name of person obtaining consent Signature of person obtaining consent

8.5. Peer Assessment questionnaire

ID: **Date:**

Dear student teacher: This form was designed to assess the quality of the explanations you are observing. Please assess it according to the scale, and justify your evaluation answering the questions. Number one means you consider the explanation had not good quality and number six means you think the explanation had a very good quality. If you are not clear in one question, you can leave it blank. Use as much space you need, using one block for each explanation.

1. About the explanation given about the **difference between mass and weight**. How do you assess the quality of the explanation for these children?

	1	**2**	**3**	**4**	**5**	**6**	
It was a very bad quality explanation							**It was a very good quality explanation**

Why?
Why do you think he/she took these pedagogical decisions? What would you do in the same situation? Why?

2. About the explanation given of the **measurement procedure of water volume**. How do you assess the quality of the explanation for these children?

	1	**2**	**3**	**4**	**5**	**6**	
It was a very bad quality explanation							**It was a very good quality explanation**

Why?
Why do you think he/she took these pedagogical decisions? What would you do in the same situation? Why?

3. Do you have any other comment about this microteaching episode? Please write it down.

8.6. Process evaluation

Sessions assessment form

Dear student: This form assesses the quality of the sessions done and facilitator's performance, in order to know your opinion and improve any aspect that could be better. Please, complete the form truthfully. The survey is confidential. If you are not clear in one question, you can leave it blank.

Instructions: Please mark the option that represents you better in front of each sentence. Number one means you totally disagree with the sentence and number six means you totally agree with it.

Sessions Identification			
Session Nº1. Date:		*Objective: this will be completed by the facilitator*	
Session Nº2. Date:		*Objective: this will be completed by the facilitator*	
Session Nº3. Date:		*Objective: this will be completed by the facilitator*	
Session Nº4. Date:		*Objective: this will be completed by the facilitator*	
University:		Facilitator:	*this will be completed by the facilitator*

	1 I totally disagree	2 I partly disagree	3 I disagree more than I agree	4 I agree more than I disagree	5 I partly agree	6 I totally agree
Session features						
The concepts and themes that were treated in the past four sessions were useful to my future work as a science teacher.						
The methodology used in the sessions was appropriate in relation to the objectives.						
The use of time during the sessions was efficient.						
The emotional climate during the sessions was appropriate to encourage students' participation.						
The interaction between teacher and students in the sessions was appropriate to reach the objectives.						
The interaction between students in the sessions was appropriate to reach the objectives.						
The sessions' objectives were reached.						
Facilitator features						
Her knowledge about content was enough to conduct the session following its aims.						
She was close enough to the students to encourage their interaction and asking questions.						
Her questions to students facilitated students' learning and reflection.						
Write here any other comment you want about how to improve the sessions, the objectives, etc.						

8.7. Focus group and interview question guide

These were the questions that facilitator covered in a semi-structured discussion at the end of the intervention and the questions used in the follow up study for individual interviews.

If other questions emerged in the meeting they were permitted if they did not take the student teacher too far from the core questions. The order was a suggestion. The facilitator could choose to follow the most adequate order for the meeting flow. However, all the questions were covered.

Focus group questions:

1. What did you expect to learn in the sessions?
2. Did you feel comfortable doing the microteaching and receiving feedback from a peer? Why?
3. Do you think was useful doing the microteaching and receiving feedback from a peer? Why?
4. (If they say yes) In what aspects was it useful? Why?
5. Do you think this method of teaching should be used with other students? Why?
6. What were the most important things that you learn in these sessions?
7. Do you think what you learned will be sustained over time – will you remember it? Why?
8. Is there any other aspect you would like to comment on about the sessions or your learning?

Individual interviews questions:

1. What do you remember from the Peer Assessment Seminar?
2. Which of these elements were the most important for you? Why?
3. Do you remember which were your strengths and weaknesses when explaining scientific concepts to your peers?
4. How are these practical aspects now when you face real teaching?
5. Do you think the skills you develop in the Peer Assessment Seminar are remaining stable, increasing or decreasing in the context of real teaching? Why?
6. What are the facilitators to transfer your skills to explain scientific concepts into real teaching?
7. What are the obstacles to transfer your skills to explain scientific concepts into real teaching?
8. Is there any other aspect you would like to comment?

8.8. Rubric to assess the quality of explanations

Criteria/ Level	Not achieved (0)	Half achieved (1)	Achieved (2)
1. Clarity	The teacher does not fulfil any of the elements or the explanation is confusing, vague or tautological	The teacher describes the concept first and then gives it a name or definition **or** uses an understandable language for the students	The teacher describes the concept first and then gives it a name or definition **and** uses an understandable language for the students
2. Coherence and cohesion	The teacher does not fulfil any of the elements	Each part of the explanation shows a relation (i.e. cause-consequence, inclusion-exclusion, differentiation, similarity) with the following part **or** The explanation presents strong cohesive elements	Each part of the explanation shows a relation (i.e. cause-consequence, inclusion-exclusion, differentiation, similarity) with the following part **and** The explanation presents strong cohesive elements
3. Sequence	The teacher does not fulfil any of the elements	The teacher explains the concept in a progressive sequence **or** The main ideas presented scaffold the concept construction	The teacher explains the concept in a progressive sequence **and** The main ideas presented scaffold the concept construction
4. Accuracy	The teacher explains the concept with inaccuracies that drive into a conceptual mistake **or** The explanation contains a conceptual mistake	The teacher explains the concept with some inaccuracies that make the concept vague but they do not imply a conceptual mistake	The teacher explains the concept without any conceptual mistake or with some generalizations necessary for learning
5. Sufficiency	The explanation presents **mainly aspects that do not contribute** to the concept construction	The explanation presents **some** aspects that contribute to the to the concept construction	The explanation presents **the main** aspects that contribute to the concept construction
6. Connection with students' experience	The teacher does not fulfil any of the elements	The teacher identifies students' prior ideas or mentions students' everyday life aspects related to the concept **without** connecting it explicitly with the explanation (i.e. integrating it, confronting it, etc.)	The teacher identifies students' prior ideas or mentions students' everyday life aspects related to the concept **and** connects it explicitly with the explanation (i.e. integrating it, confronting it etc.)

(cont.)

Criteria/ level	Not achieved (0)	Half achieved (1)	Achieved (2)
7. Metaphor, analogy, simulation or model usage	The teacher does not use rightly a metaphor, analogy, simulation or model to explain	The teacher uses rightly a metaphor, analogy, simulation or model to explain **without** mentioning the concept features that are present in the metaphor, analogy, simulation or model	The teacher uses rightly a metaphor, analogy, simulation or model to explain **and** mentions the concept features that are present in the metaphor, analogy, simulation or model
8. Example, demonstration, experiment, graphic or image usage	The teacher does not fulfil any of the elements	The teacher uses an example, demonstration, graphic or image to complement the explanation **but** The teacher does not illustrate, clarify or highlight a concept feature through it	The teacher uses an example, demonstration, graphic or image to complement the explanation **And** The teacher illustrates, clarifies or highlights a concept feature through it
9. Gesture and voice usage	The teacher does not fulfil any of the elements	The teacher uses appropriately her gestures to complement the explanation **or** The teacher uses appropriately the voices to highlight some aspects of the explanation	The teacher uses appropriately her gestures to complement the explanation **and** The teacher uses appropriately the voice to highlight some aspects of the explanation
10. Misconception illustration	The teacher does not mention any common misconception in the understanding of the concept	The teacher mentions a common misconception in the understanding of the concept **without** mentioning how students can avoid it	The teacher mentions a common misconception in the understanding of the concept **and** mentions how students can avoid it

Note: this rubric indicates “scientific concept” as a general denomination for scientific phenomena, scientific principle, scientific concept, scientific postulate, scientific situation or context.

8.9. Subcategories description of student teachers' conceptions

1. General Aspects (GA): This category grouped together the aspects that teachers mentioned which were not related with the way the teacher delivered a conceptual explanation. However, they might be related with other general aspects of teaching.

1.1 Lesson preparation: It refers to the existence (or not) of previous preparation that the teacher did to create the lesson presented in the microteaching episode.

> *'You prepared the lesson, you brought a slide, you brought little figures, and then I see you were prepared in advance for this activity.' (U1, T1:1)*

1.2 Formal clothes: It is a comment about how the teacher dressed for the microteaching episode.

> *You were showing a good looking (for the occasion)! (U1, T6:5)*

1.3 Voice, diction and rhythm: It groups ideas about the use of voice (voice projection), the pronunciation in the teacher's speech and the rhythm (voice inflection, pauses, highs and lows).

> *You had a good disposition with the class, your voice tone, but I think it was very low. (U1, T5:150)*

> *And well, the fluency when you are speaking, that is good, because before you spoke very fast. Now it is good, it is paused, as it should be. (U1, T4:217)*

1.4 Values in teaching: It refers to how the teacher associated societal values with the science content of the lesson.

> *He included a very important topic, the societal value topic, and sometimes we leave it out, we go to the context, to the content and setting aside why is this useful for us? Then we have to always rescue the value behind. And usually it is taken at the end and not at the beginning, not during the lesson, or constantly as it should be. This is very important because we want to give societal values to the pupils, then it is good to take it from the beginning, and it is also important at the end. (U1, T1:64)*

1.5 Content vs. time relation: These are comments related to the time management during the explanation, generally referred to how the teacher handled the little time in the microteaching episode related to the content the teacher reviewed.

But here you had the opportunity to choose only one concept... but you included too many concepts in the time we had. (U3, T16:140-142)

1.6 Excess of content: It refers only to the excessive quantity of content given in the microteaching episode without linking it with the time available.

You gave so many contents that sometimes I could not process them in a good way, because it was one concept followed by the other. When one of them was more slowly I could think "ok, this concept... I could link it with this other thing", but when there are that many concepts it is complicated to process the information. (U1. T2:125)

1.7 Clarity of handwriting: It groups comments given about how clear (or not clear) the teacher's handwriting was and the consequences that it could bring to the pupils.

The handwriting! Because the pupils are going to try to do the same letter the teacher has. Then if the teacher writes "a" "c" with an "l" in between, then they say "ka, does it say la?" For the teacher it is a "r" or a "l"... then they are going to pay more attention to write what the teacher is writing in the whiteboard than to what the teacher is explaining. (U3, T16, 35).

1.8 Image and speech synchrony: They are observations related to the importance of speaking and showing images or any graphic resource at the same time.

I try to link what I am saying with one part of the image. Because I could be speaking to the air, but I could be referring to the image as a total. (U3, 13:211)

1.9 Importance of contents: It refers to ideas about how important the content is.

I think these concepts are very strong, they should be base... (U2, T11:11)

1.10 Movement in the classroom: It summarises observations about the ways in which the teacher moved around the classroom (if the teacher stayed static or used the space available).

Maybe she lacked of displacement around the classroom, she stayed static there... I think the displacement. (U3, T17:162-166)

1.11 ICTs usage: It refers to the presence or absence of PowerPoint™ (Microsoft, 2010) presentations or other multimedia resources.

I mean, I liked the PowerPoint presentation. But it had too much text. (U3, T14: 139)

1.12 Emphasis on note-taking: These are comments referring to teachers' emphasis in the importance of pupils' taking notes process during the explanation.

The good point is he asked the pupils to take notes about the explanation. (U1, T5: 11)

1.13 Whiteboard layout: It summarises comments about how accessible the layout of teacher ideas on the whiteboard (including the design of schema, the drawings, etc.) is for pupils to follow, and the interaction between the teacher and the written ideas.

Then while I am doing more schemas, the whiteboard layout is becoming messier. Then I think I should look after this aspect, doing a table and try to do not go over or away from the table, or dividing the whiteboard in two… (U1, T2:190)

1.14 Lesson title: It refers to the absence or presence of a written title in the lesson to orientate pupils.

The lesson doesn't have a title. You mentioned it but it doesn't have title. (U1, T4:198)

1.15 Self-confidence: It groups the observations about how confident the teacher appeared or looked like at the moment of explaining.

I saw her nervous, and at the beginning like not very sure what she was saying. Well, after she felt more comfortable and she was confident, but at the beginning I think her voice was shaking and that showed me nervousness. (U2, T12:106)

1.16 Difficulties overcoming: It refers to how the teacher, during the microteaching episodes or personal teaching experience, is dealing with unexpected events and overcoming them in order to continue with the explanation.

Although the PowerPoint presentation did not work in a moment, she overcame it quickly, and it did not affect the explanation. (U3, T7:123)

2. Pedagogical Knowledge Aspects (PKA): Under this category all the comments about relevant issues in teaching through conceptual explanations were grouped together. It shows pedagogical knowledge present in the teachers' mind at the moment of assessing the quality of explanations. However, the comments in this category were not linking the pedagogical knowledge with the content the teacher was addressing in the explanation.

2.1 Resources availability: It refers to the presence or absence of teaching resources.

I think the figures as a resource it was a good input to explain. (U1, T4:2)

2.2 Adequacy for pupils' characteristics: It shows how adequate the explanation was for certain pupils' characteristics (interest in a topic, lack of concentration, etc.).

In general I like the attitude that T1 has, it is like challenging. But it depends on the pupils. In my case the pupils do not feel motivated for learning, then they do not have a goal, a challenge, nothing, then they are different contexts. Because they are children who feel upset with that and the teacher needs to be aware of that. It depends on the context, because if you know your pupils, you could challenge them. If you do not know them and you come like that, it is complicated. (U1, T6:85)

2.3 Adequacy for teaching phase: It groups together comments about whether the explanation was adequate or not considering the teaching phase the teacher performed.

I think the figures may be used after, to classify… because they were in the introduction of new points of view phase, he tried to use them maybe vaguely, but maybe in a future structuration phase I think they could be very useful. (U2, T2:8)

2.4 Activity goals: Regarding the learning intention or goal of the activities the teacher presented in the microteaching episode.

Because if you start like that, pupils may be diverted to either side, and they could be diverted from the goal, then they must always know that there is an objective to achieve, and they will focus on that. (U3, T15:15)

2.5 Content complexity v pupils' age: It summarises observations about how complex the content was, or how deep was the content reviewed, related with the pupils' age or grade.

Yes, that content is given in 8th grade, not in 6th. With the topic you saw there it is seen in 8th. The topic is given in 6th but basic. But when you start speaking about the process, it should be seen in 8th. (U1, T3:98,100)

2.6 Pedagogical language usage: It regards the use of language that is desirable during teaching (formal, clear, and familiar for pupils).

I like the vocabulary T4 uses. I think she is very "teacher" when she speaks… (U1, T2:28)

And also in the questions I think you were very basic in the use of language, like "what you want say". When you express yourself you do the same. You say some words in between, some rare things, and that is taking my attention some times. It is more about the way you speak. You need more what T4 is doing. (U1, T2:156, 160)

2.7 Resource usage: It refers to what the teacher did with the resources in the microteaching episode, the decisions they took during the lesson, the interaction of the pupils with the resource, etc.

And it is true that is important, because the pupils ask "How do I distribute my space, so I put the sheet like this, where I should leave more space?" And after just 10 or 15 minutes you used the table. Then you should make the table just when you are going to use it. (U1, T4:229)

2.8 Pupils' participation: It regards how the teacher encourages pupils to participate or give their opinion during the explanation.

I think it was good to ask the pupils to go to the front, making them participate. (U2, T14:79)

2.9 Question type and usage: It refers to the type of questions (open, closed) the teacher asked, and what was the function or use of the question.

And then, the question you did I don't know if is the most successful as initial question for the exploration, it should be a question more opened, and your question was totally closed. (U1, T4:101)

2.10 Resource characteristics: It groups together several comments about different characteristics the resources the teachers used, like the colour, the tidiness, the identification of parts, how understandable it was, etc.

The material she used had good information, it was the same he was supporting on, and the drawings presented on the whiteboard were exemplars and clear about the topic being addressed at the time. (U3, T16:134)

2.11 Pre-concepts gathering: It refers to the inclusion or not within the microteaching episode of any technique to gather pupils' pre-concepts.

I think it was missed, maybe writing the brainstorming the pupils were giving, to be able to take them again after... From I could have seen, the children are also going to see this, but maybe some of the ideas that were not there you could write them in the whiteboard, maybe. (U3, T16:178-180)

2.12 Collective construction of knowledge: It groups together observations referring whether the teacher shared the concept construction with the pupils or encouraged pupils to work the concepts in groups or teams.

Because in every moment T2 constructed. He constructed drawings, he constructed the definitions, the schema, and everything was through the construction of the pupils and himself. (U1, T4:68)

2.13 Questions delivery: It summarised ideas about how the teacher posed the questions to the class during the microteaching episode, i.e. to one student only, to the whole class, a random student, etc.

> *Maybe it was not clear if the question was for the pupils of for herself (U1, T3:46).*

2.14 Answers' management: It refers to how the teacher dealt with the answers to their questions during the explanation.

> *He took the pupils pre concepts during the questions, but he did not take it in the answers, when you said a wrong answer he did not move it, he just said no. (U1, T5:11)*

2.15 Schemas title: It referred to ideas regarding if the schemas shown or drawn by the teacher have a title or not indicating what was the schema presenting.

> *And the drawing in the schema also did not have a title, in the summary. It would have been better if it would have title, because you made mistake (U1, T4:211)*

2.16 Examples usage: It grouped together the comments about the quantity of examples during teaching, and if they were close to the pupils' experiences.

> *The positive point is I believe she gave several examples about what she was doing, each one of the aspects. (U2, T7: 99)*

2.17 Activity type: It refers to the type of activities present in the explanation.

> *Because what I missed an activity... maybe a group activity, I do not know. (U2, T8:120)*

2.18 Gesture usage: It grouped together ideas related to how the teacher uses gestures and certain type of non-verbal communication (such as eye contact) at the moment of explaining.

> *I think it's good that she uses the gestures, because they take the attention. (U2, T14:198,202)*
>
> *But I would change a little more the expression that I have towards others, to look at them, messing around with that, the feedback. (U3, T18:147)*

2.19 Connection with scholar texts: It was a comment connecting the way the teacher explained to the regular illustrations appearing in scholar texts.

> *The thing is most of the school texts now come with that, with that schema. (U2, T8:8)*

3. Pedagogical Content Knowledge Aspects (PCKA): This category refers to all teachers' ideas that teachers had to assess the explanations that showed applied pedagogical knowledge to the content. These ideas mean the teachers were thinking of the ways to teach more effectively a specific piece of content or how learning could be more meaningful, which implies a deeper applied and more flexible knowledge.

3.1 Resource adequacy for content or goal: It refers to how adequate the resource was considering the content or the goal the teacher wanted to get across. It includes the resources' appropriateness for the teaching purposes.

> *Sure, the images were presented according to the content, nothing was disconnected from it. The images he presented were totally showing the differences he wanted to present, both the positive and the negative aspects of the same force. (U3, T16: 12)*

3.2 Nature of science: It grouped together comments about what kind of characteristic of the nature of science was transmitted implicitly or explicitly during the explanation, i.e. Science as: constructed, flexible, representational or fixed, static, transmissible.

> *Then, I consider that always it is a good opportunity to introduce the diversity issue. I mean, we are different people, we are not equal, then the menstrual cycle may be 21, 28, 35 or even more days, then the self-care there... Yes I think that at all times, for example I saw this with 8th grade and I said, girls, you are at this moment ... I remember I had to do something similar, and I brought the issue of diversity. (U1, T6:45,47)*

3.3 Analogy accuracy: It refers to the accuracy of analogies or comparison the teacher presented during the explanation to provide a better understanding of content.

> *The same analogy you must do it with caution, because this about clean water and dirty water... because it's like you have clean and dirty blood. But instead of dirty blood, because it is not that it is dirty, but it has waste. You must clarify some aspects that can be misunderstood with the analogy. (U1, T2:112)*

3.4 Quality of resource: It includes comments about different aspects or characteristics the resources had which could make them being better or worse in terms of the quality of the explanation.

> *And regarding your schema, I think in the vertical plane it was fine, but in the horizontal, you missed some things, such as putting there "hormone, target organ, action", etc. (U1, T2:32)*

3.5 Pupils' ideas integration: It refers to whether the teacher integrated the pupils' ideas about the content into the explanation or not.

> *She gathered the children's prior ideas. Like what was a plastic, which were the plastic elements we had inside the room, glass, wood, or what were the elements present in the classroom. And*

she asked questions for the full class, and she was incorporating... the answers were fully incorporated in the subject she was working. (U3, T16:15)

3.6 Mistakes management: It refers to what the teacher did when pupils' mistakes or content misconceptions appeared during the explanation.

Once, T3 said something, a concept when he said ... the difference between chain and net. The answer the teacher gave was in a way that we started making fun of T3. During the lesson this is complex, because you have to try although he is saying something 100% wrong, you must orient him the correct answer or to something correct. (U1, T4:71)

3.7 Content contextualization: It summarises observations about how the teacher gave a context during the explanation.

She did not contextualize, there was no relation to the context. She only generated a link to the context within the subject, but it was conceptual, not connected with pupils. She had opportunities to produce some contexts, because we were so out of context to be able to understand. (U1, T5: 1)

3.8 Connection with other contents: It refers to the connection the teacher established between the content the teacher was working with and other areas, contents or ideas.

I could have gone on talking of digestive system, endocrine, urinary, but my question was referring to what was nutrition. I should have made the connection. (U1, T5:127)

3.9 Questions' specificity to the content or goal: It grouped comments regarding how specific the teachers' questions were in relation to the content or the goal the teacher wanted to achieve with the explanation.

For example T6 asked "what they do". And pupils answered something, and she said "no, but in the organ". Then "they do" points to the function, but in her speech she was expecting something specific. (U1, t2:19)

3.10 Example adequacy: It refers to how adequate the examples were to illustrate the content the teacher was working with.

I think you lacked of drawing something to exemplify where the prostate is in this structure. And it is not, there it was missing ... (U1, T2:38)

3.11 Connection between explanation and goal: It refers to how adequate the complete explanation was considering the goal the teacher wanted to achieve.

The goal definition talks about the flow of energy and the matter, but finally you based the explanation itself on the definition of what would be the ecosystem. (U1, T4:69)

3.12 Explanation sufficiency for the content: It refers to whether the explanation contained enough information or not in relation to the content, or there were missing or unnecessary parts for the explanation.

> *I think you just lacked a bit of information about the concept, for example in the water cycle topic, etc... (U2, T8: 11)*

3.13 Didactic transposition: It groups together all the comments about how to transform complex scientific concepts in ways which are accessible for young learners. Most of the comments were suggestions to make content more understandable.

> *Having placed a trophic chain, with producer and consumer, and decomposer is outside of those elements, if you had put the decomposer down with an arrow pointing above. Because the decomposer eats the plant, the rabbit and the fox. And then, you go to the trophic net. It's like a good bridge. (U1, T6:65)*

3.14 Pupils' ideas transformation: It refers to how to transform the pupils' ideas about the content during the explanation or the importance of doing it.

> *For example I told you "the glands and the messengers". And you wrote "messenger" then you took only what was useful for you. The idea is you could write down verbatim as we said, and transform them and change them according to their utility. (U1, T4:168)*

3.15 Explanation sequence: It grouped all the assessment comments related to how the explanation was structured in terms of sequence, in a progression, order of the elements, following a thematic line, or comments about which part of the explanation could be given first or after the others to achieve the content understanding.

> *I think T6 should have in this explanation a more structured chronological order, because if you ask the question, it should go after the goal, because the goal goes first, then the question. (U1, T2:235)*

3.16 Thinking skills: It refers to how the conceptual explanation the teacher gave could imply development of pupils' thinking skills, or the importance to do this.

> *What I liked in this explanation of T11 is that she always tries to keep children interpreting through the images. And it is good to make children think a little bit, because T11 could have said "here this and that happen in the picture", but not. (U2, T10:98)*

4. Knowledge Aspects (KA): This category includes comments referring to the science knowledge the teacher handled. They were given as comments about the accuracy of teacher's explanation in terms of science content expressed in scientific terms, algorithms or processes.

4.1 Scientific terms accuracy: It refers to the comments teacher gave about how correct the scientific terms used in the explanation were. It includes ideas about accuracy in specific content or general comments about the use of scientific words.

> *But the protons are not expanded, they are not moving from one atom to another ... Except in a circuit. In the circuit that you do, like here, they are protons moving, but they are minimal. The electrons are mostly moving. (U2, T11:28-30)*

4.2 Up-to-date knowledge: It groups together comments about the need of knowledge actualization, due to scientific terms which have recently changed. This implies teacher have to look for the data or review the concepts he is presenting. It is important to notice these ideas were not meaning that the teacher's concepts were wrong.

> *I would say to T4 also to be updated, because "fallopian tubes" are now called oviducts. (U1, T2:32)*

4.3 Processes accuracy: It refers to the accuracy in the processes the teacher is describing or the scientific phenomenon he is presenting. The focus is in the phases of the process and the accuracy in how it works.

> *I think the minutes need to be transformed into seconds. As I knew at least, they are converted to seconds. (U2, T11:2)*

4.4 Algorithms accuracy: It refers to the accuracy in step-by-step procedures for calculations, such as equations or mathematical procedures that are being taught as part of science content.

> *I think that equation is wrong! (U2, T8:1)*

8.10. Map of codes Focus groups

Category	ID	Subcategory	Frequency	Example
1. Phenomena	1.1	Changing theories	7	*And everybody points to the conceptual change, and the conceptual change is not black or white, our ideas were changing gradually*
2. Causal conditions	2.1	Confrontation theory-practice	3	*And because each of us says we follow the constructivism but in the practice, but when you stand in front of the class it is different*
2. Causal conditions	2.2	Change of analysis focus	12	*I initially looked at what happened to the children, and now I focused more on what happened to the teacher. Then I was completely out of focus in the first activity. And about the teacher ... now I focus on what she decided to do, or what she did wrong, etc.*
2. Causal conditions	2.3	Reflect or projection	7	*It's different now because we put in the other's place ... It helps us put ourselves in the place of the other, what he wanted to accomplish or what was he expecting with the thing he did, and without evaluating ourselves it would not have been possible*
2. Causal conditions	2.4	Reflection on practice	6	*I would consider that here teacher reflection process is much more valuable that how he made the lesson*
2. Causal conditions	2.5	Visibility of weaknesses and strengths	3	*I think it is very important to know your own weaknesses. Because knowing what we're missing, we can work to improve it while maintaining the strengths*
3. Context	3.1	Respectful constructive critique	5	*What I valued the most was the critique in a respectful framework*
3. Context	3.2	Peer assessment with role interchange	5	*I think so, it's a good way to learn, evaluating yourself and evaluating others*
4. Intervening conditions	4.1	Knowing new criteria	4	*So I think having the assessment parameters it helps a lot, because I know what my classmates are looking for: this, this and this. Then that is useful to outline the lesson*
4. Intervening conditions	4.2	Systematic evaluation of practice	3	*This would benefit also in our own placement centres, because from my point of view they assess one session only, and sometimes it does not reflect the entire process of teaching*
4. Intervening conditions	4.3	Empathetic facilitator	1	*It is noticeable that she not only does it because she has to complete the task or because she must cover certain things, but she does so in a human way and this is also much appreciated.*

Category	ID	Subcategory	Frequency	Example
5. Strategies	5.1	Self-regulation and improvement	7	*Then you realize, you appropriate this mistake, and you are able to modify it and not do it again. I am more cautious now.*
5. Strategies	5.2	Negotiation of meaning	3	*But the main thing here is that the criteria were made from group consensus. And I think the consensus on any assessment is important, and in this aspect we were fine*
6. Consequences	6.1	Changing practices	6	*I see changes in how I do it, how to evaluate and how I'm going to perform doing a lesson and self-evaluating*
6. Consequences	6.2	Transference of PA to school context	4	*I think it adds a lot in terms of our future employment skills. And this programme could continue in the labour context with our peers, to be able to improve. It could be done. If you could agree with your colleagues to assess the practice and also with self-assessment, it would improve the quality of our work, the teaching process with children, etc.*
6. Consequences	6.3	Sharing learning with other student teachers	4	*So I think this should be done from the first year of training, subject by subject, it does not matter what it is covering in terms of content, it can be geometry, whatever, but working with the explanations it would help every subject. For example there were several things that we realized that we did not handle, and we had to prepare them for the lesson, to study and do it. So I think it's super important to have opportunities like this*
6. Consequences	6.4	Openness to criticism	1	*And the disposition changes also. Now I am more open to criticism. I think this is the basis of all. More open to receive them.*

8.11. Map of codes Interviews

Category	ID	Subcategory	Frequency	Example
1. Facilitators	1.1	Guide teacher	4	*My guide teacher motivates me, I think she is one of the only good teachers there are in the school, because the others do not have a good academic level, she is one of the few that have it ... She is a very good teacher, she knows how to teach and she handles the content*
1. Facilitators	1.2	School support or flexibility	3	*There is a lot of flexibility in this school to do the type of lesson I want*
1. Facilitators	1.3	Criteria construction and critique	3	*I think our ability to create an instrument was very important because helps us to improve our own practices. Then, from what we have created, we correct ourselves now.*
2. Mediators	2.1	Classroom climate	5	*There is also the factor of classroom climate. Here the pupils are not so messy, but it is not the ideal context. There are always pupils that disturb the class, which are putting others pupils off*
2. Mediators	2.2	School resources	4	*Also, the resources are given to you at this school, although they are basic. We have computers, we have materials, have back garden, theatre for a play, etc.*
2. Mediators	2.3	Teachers' CK and knowledge of pupils	4	*I advance in certain areas, depending on the subject matter, I mean depending on how I handle the content.*
3. Obstacles	3.1	Lack of time for planning and reflection	2	*I would like to have more time to prepare lessons, because time is a crucial determining factor. I dedicate the weekend to do it, in between that I have to have time for family life, and now for example I am taking paper work to be done while I teach the other lesson*
3. Obstacles	3.2	Low school interest in science education	2	*In this school, science as an everyday life thing does not exist*
3. Obstacles	3.3	Pupils' lack of participation	2	*Well, in this case I could easily explain and explain generating a monologue. But when you ask questions to the students and you make students participate, you notice here students do not participate when I ask them.*
3. Obstacles	3.4	Pupils' lack of science knowledge	3	*The pupils had problems with previous teachers, and then they have some content deficiencies*
3. Obstacles	3.5	Weak and disconnected ITE	2	*But I think that lessons we received in the university were planned considering that we will have an ideal class, where students' skills are high, where the classroom climate is good, where we are not considering the problems that students are exposed*

8.12. Authorization for use of image

Permit for using and modifying the image presented in Chapter 5, Figure 20 was obtained from Graciela Lobos González, Physics teacher at Pumahue School in Temuco, Chile. It was obtained via email, as shown in the following sequence of messages (in Spanish).

RE: Contacto de Valeria Cabello (v.m.cabellogonzalez@dundee.ac.uk) en "Sala de Física"
Sent:11 December 2012 15:14
To: Graciela Lobos González [globoscl@yahoo.com]

Muchas gracias Graciela!
Exactamente me fue muy complicado encontrar una buena imagen, tu sitio fue un acierto. De todos modos voy a modificar la imagen pues le voy a poner texto para indicar lo que representa cada parte. Muchas gracias por tu permiso. De todas maneras voy a citar tu blog en mi tesis para clarificar desde donde obtuve la imagen.
Gracias nuevamente,
Valeria

From: Graciela Lobos González [globoscl@yahoo.com]
Sent: 11 December 2012 15:08
To: Valeria Cabello Gonzalez
Subject: Re: Nuevo mensaje de contacto de Valeria Cabello(v.m.cabellogonzalez@dundee.ac.uk) en "Sala de Física"

Hola Valeria
No hay problema en el uso de la imagen ya que yo misma la descargué desde internet. No es fácil encontrar imágenes buenas relacionadas con conceptos físicos, a pesar de que hay muchas.
Hasta pronto
Graciela Lobos González
Temuco

De: Bligoo <noreply@bligoo.com>
Para: globoscl@yahoo.com
Enviado: Jueves, 6 de diciembre, 2012 8:13 P.M.
Asunto: contacto de Valeria Cabello(v.m.cabellogonzalez@dundee.ac.uk) en "Sala de Física

Hola Graciela. Soy Valeria Cabello, estudiante de doctorado chilena. Quisiera saber si me darías autorización para usar la imagen de refracción de luz que aparece en tu página web para ilustrar un punto de mi investigación. Yo la modificaría agregándole texto. Me parece una excelente imagen. Ojala me puedas responder pronto a mi correo electrónico.
Muchas gracias de antemano,
Valeria